

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 11, 2003, 19:43:59 ; Search time 8.74286 Seconds
(without alignments)
2096.859 Million cell updates/sec

Title: US-09-497-967-6

Perfect score: 2342

Sequence: 1 MKYNILLIILISLFINELRA.....STTFAKFLISLIFISFYLL 442

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	213.5	9.1	713	1	TSA4_GIALA
2	202.5	8.6	3718	1	LMA5_MOUSE
3	198	8.5	3084	1	LMA1_MOUSE
4	191	8.2	3110	1	LMA2_HUMAN
5	189	8.1	5376	1	ZAN_MOUSE
6	187	8.0	328	1	C170_GIALA
7	182	7.8	667	1	TS11_GIALA
8	182	7.8	3695	1	LMA5_HUMAN
9	181.5	7.7	3075	1	LMA1_HUMAN
10	181	7.7	687	1	VS41_GIALA
11	178	7.6	3106	1	LMA2_MOUSE
12	177.5	7.6	1607	1	LMG1_MOUSE
13	175	7.5	1587	1	LMG3_HUMAN
14	173	7.4	1790	1	LMG1_DROME
15	172.5	7.4	1581	1	LMG3_MOUSE
16	172	7.3	913	1	PKS5_HUMAN
17	171	7.3	1576	1	YK83_CAEEL
18	169.5	7.2	1169	1	YK82_YEAST
19	169	7.2	1877	1	PKK5_MOUSE
20	168.5	7.2	1696	1	PKK5_BRACL
21	168.5	7.1	1557	1	LMG1_CAEEL
22	165.5	7.1	1680	1	FUR2_DROME
23	165.5	7.1	1700	1	BAR3_CHITE
24	165.5	7.1	3712	1	LMA_DROME
25	162.5	6.9	2704	1	G168_PAPPR
26	160	6.8	1609	1	LMG1_HUMAN
27	158.5	6.8	600	1	SP96_DICDI
28	158	6.7	1246	1	YMW2_CAEEL
29	157	6.7	790	1	ANP_NOTCO
30	156.5	6.7	2713	1	G156_PAPPR
31	156	6.7	1192	1	LMG2_MOUSE
32	154.5	6.6	1786	1	LMG1_MOUSE
33	154.5	6.6	2556	1	NTC1_HUMAN

34	152.5	6.5	2471	1	NTC2_RAT
35	151.5	6.5	3224	1	RBP2_HUMAN
36	150.5	6.4	1193	1	LMG2_HUMAN
37	149	6.4	2437	1	NTC1_BRASE
38	147.5	6.3	2470	1	NTC2_MOUSE
39	146.5	6.3	1799	1	LMB2_MOUSE
40	146	6.2	1797	1	VGLX_HSVB
41	145.5	6.2	1786	1	LMB1_HUMAN
42	144.5	6.2	2531	1	NTC1_MOUSE
43	143.5	6.1	2569	1	LMA3_MOUSE
44	143	6.1	2524	1	NOTC_XENLA
45	142.5	6.1	1877	1	PKK5_RAT

O9q430 rattus norv
P49792 homo sapien
Q13753 homo sapien
P46530 brachydanio
O35516 mus musculu
Q61292 mus musculu
P28968 equine herp
P07942 homo sapien
Q01705 mus musculu
Q61789 mus musculu
P21783 xenopus lae
P41413 rattus norv

ALIGNMENTS

RESULT 1

TSA4_GIALA
ID TSA4_GIALA STANDARD; PRT; 713 AA.
AC P21849;
DT 01-MAY-1991 (Rel. 18, Created)
DT 01-MAY-1991 (Rel. 18, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Major surface-labeled trophozoite antigen 417 precursor.
GN TSA 417.
OS Giardia lamblia (Giardia intestinalis).
OC Eukaryota; Diplomonadida; Hexamitidae; Giardiinae; Giardia.
OX NCBI_TaxID=5741;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=ATCC 30957 / WB;
RX MEDLINE=90280395; PubMed=2352929;
RA Gillin F.D., Hagblom P., Harwood J., Alely S.B., Reiner D.S.,
RA McCaffery M., So M., Guiney D.G.;
RT "Isolation and expression of the gene for a major surface protein of
Giardia lamblia.";
RL Giardia lamblia.;
RL Proc. Natl. Acad. Sci. U.S.A. 87:4463-4467(1990).
RN [2]
RP SEQUENCE OF 480-620 FROM N.A.
RX STRAIN=AD-1;
RX MEDLINE=93314970; PubMed=8325510;
RA Ey P.L., Mayrhofer G.;
RT "Two genes encoding homologous 70-kDa surface proteins are present
within individual trophozoites of the binucleate protozoan parasite
Giardia intestinalis.";
RL Gene 129:257-262(1993).
CC -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN. ANCHORED TO THE
PLASMA MEMBRANE.
CC -!- DOMAIN: CONTAINS 29 REPEATS OF THE CXXC MOTIF.
CC -!- SIMILARITY: BELONGS TO THE GIARDIA VARIANT SURFACE PROTEIN FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL Outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC
CC EMBL; M33641; AAA02688.1; -;
CC EMBL; M97488; AAA02581.1; -;
CC PIR; A35502; A35502.
CC InterPro; IPR000561; EGF-like.
CC InterPro; IPR002174; Furin-like.
CC InterPro; IPR005127; Giardia_VSP.
CC Pfam; PF03302; VSP; 2.
CC SMART; SM00181; EGF; 1.
CC SMART; SM00001; EGF-like; 1.
CC SMART; SM00261; FU; 3.
KW Signal; Antigen; Glycoprotein; Transmembrane; Repeat.
FT SIGNAL 1 17
FT CHAIN 18 713 MAJOR SURFACE-LABELLED TROPHOZOITE

Thu Feb 20 11:30:52 2003

FT	DOMAIN	18	679	ANTIGEN 417.	
FT	TRANSMEM	680	708	EXTRACELLULAR (POTENTIAL).	
FT	DOMAIN	709	713	POTENTIAL.	
FT	CARBOHYD	289		CYTOLASMIC (POTENTIAL).	
FT	CARBOHYD	676		N-LINKED (GLCNAC. . .) (POTENTIAL).	
FT	VARIANT	582		N-LINKED (GLCNAC. . .) (POTENTIAL).	
FT	VARIANT	606		A -> T (IN STRAIN ADELAIDE-1).	
FT	SEQUENCE	713 AA;	72510 MW;	A -> S (IN STRAIN ADELAIDE-1).	
QY	Query Match	9.18;	Score 213.5;	DB 1;	Length 713;
QY	Best Local Similarity	23.08;	Pred. No. 1.3e-07;		
QY	Matches 127;	Conservative 42;	Mismatches 187;	Indels 197;	Gaps 31;
QY	42	GTCVNCRFNYNGAAQ-----GEANGNQFPANNAARGI--CVPCQINRVSVTN	91		
Db	111	GVCTEAGYFAPVGAANTEQSVIACGDTTGTIAAGNTYKGIADCAECSAPDATAGAE	170		
QY	92	AGDLATLATQC-----STQCPTGT-----ALDD-----GV-	116		
Db	171	AGKVAT-CTKGVSKYLDNVCVDAQCNSGSTNKFVAVDDSENGNKCVCSDNLNGGVA	229		
QY	117	---TQVEDRSA--AOCVKCKPNFY-----NGSPQGE-APGVQVFAAG	154		
Db	230	NDCYSDQSKIKCTKCTDNNYLLKTTSEGTSCVQKQDKGFFPKDDSSAGNKCCLPCN	289		
QY	155	AAAAGVAAVTSCVPCQLNKNDSPA-----TAGAQAQLAT-----QC	191		
Db	290	DSTDGIA-----NCATCALVSGRGAALVTCSACTDGYKPSADKTTCEAVSNCKTPGCKAC	345		
QY	192	SNQCPGTGLD--DGVTLVNTISAFL-----SCKTSAGTA---SDCTECPGKALRYGDDG	455		
Db	346	SNEGKENEVCTCDGSLTYLTPISQIDSCAKIGNYGGATEGAKKLCRETAANCKTCDQ	405		
QY	216	--CVKCRPNFYNGG--SPQGEAPGVQVFAAGAAAGVAAVTSCVPCQINK-----NDS	266		
Db	406	GQCOACNDFYKNGDACSPCHE-----SCKTSAGTA---SDCTECPGKALRYGDDG	455		
QY	267	P-----ATAGAQAQLATQCS-----TOCPTGTAI--QDGV-----TLVFSNS	301		
Db	456	TKGTCGEGCTTGAGACACTGCLTIDGASYCECATTEYFONGVPCAPKASRAIPTCND	515		
QY	302	STQ---CSQCIANYF--FNGN-PEA---GKSQCLKCPVSKTTPAHAPGNATATQCLIT	352		
Db	516	PIQNGVCGTCADNYFMNGGCVETVPKGTVCISAPNGTGTCQAADGYKLDSGT--LT	573		
QY	353	CPAGTVLDDGTSTNFVASATECTKCSAGFF--ASKTTGFTAGTDTCTECKRLTSGATAK	410		
Db	574	CSEG-----CKEASSTDTCTCLDGYVKSASACTKCDASCEFCNGAATCKACATG-	624		
QY	411	VYAEATOKVQCAS 423			
Db	625	YKTAGSEGACTS 637			
RESULT 2					
ID	LMA5_MOUSE	STANDARD;	PRT;	3718 AA.	
AC	061001;	09JHQ6;			
DT	01-NOV-1997	(Rel. 35, Created)			
DT	15-JUN-2002	(Rel. 41, Last sequence update)			
DE	15-JUN-2002	(Rel. 41, Last annotation update)			
DE	Laminin alpha-5 chain precursor.				
GN	LMA5.				
OS	Mus musculus (Mouse).				
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.				
OX	NCBI_TaxID=10090;				
RN	[1]				
RP	SEQUENCE OF 1-92 FROM N.A.				
RA	Timpl R., Sasaki T.;				
RT	"Completion of the N-terminal sequence of the murine Laminin alpha 5				
FT	chain.";				

RL	Submitted (JUL-2000) to the EMBL/GenBank/DDBJ databases.	
RN	[2]	
RP	SEQUENCE OF 84-3718 FROM N.A.	
RC	STRAIN=CS7BL/6 X CBA; TISSUE=Lung;	
RX	MEDLINE=96081906; PubMed=7499364;	
RA	Miner J.H., Lewis R.M., Sanes J.R.;	
RT	"Molecular cloning of a novel laminin chain, alpha 5, and widespread	
RT	expression in adult mouse tissues.";	
RL	J. Biol. Chem. 270:28523-28526(1995).	
RN	[3]	
RP	REVISIONS.	
RA	Miner J.H., Lewis R.M., Sanes J.R.;	
RL	Submitted (NOV-1997) to the EMBL/GenBank/DDBJ databases.	
CC	-!- FUNCTION: Binding to cells via a high affinity receptor, laminin	
CC	is thought to mediate the attachment, migration, and organization	
CC	of cells into tissues during embryonic development by interacting	
CC	with other extracellular matrix components.	
CC	-!- FUNCTION: ALPHA-5 CHAIN MAY BE THE MAJOR LAMININ ALPHA CHAIN OF	
CC	ADULT EPITHELIAL AND/OR ENDOTHELIAL BASAL LAMINAE.	
CC	-!- SUBUNIT: Laminin is a complex glycoprotein, consisting of three	
CC	different polypeptide chains (alpha, beta, gamma), which are bound	
CC	to each other by disulfide bonds into a cross-shaped molecule	
CC	comprising one long and three short arms with globules at each	
CC	end.	
CC	-!- SUBCELLULAR LOCATION: EXTRACELLULAR; FOUND IN THE BASEMENT	
CC	MEMBRANES (MAJOR COMPONENT).	
CC	-!- TISSUE SPECIFICITY: IN ADULT, HIGH LEVELS IN HEART, LUNG, AND	
CC	KIDNEY; LOWER IN BRAIN, MUSCLE AND TESTIS; VERY LOW IN LIVER, GUT	
CC	AND SKIN. EXPRESSED IN MANY TISSUES IN EMBRYONIC DAY 11.	
CC	-!- DOMAIN: THE ALPHA-HELICAL DOMAINS I AND II ARE THOUGHT TO INTERACT	
CC	WITH OTHER LAMININ CHAINS TO FORM A COILED COIL STRUCTURE.	
CC	-!- DOMAIN: DOMAINS VI, IV AND G ARE GLOBULAR.	
CC	-!- SIMILARITY: CONTAINS 1 LAMININ N-TERMINAL DOMAIN (DOMAIN VI).	
CC	-!- SIMILARITY: CONTAINS 21.5 LAMININ EGF-LIKE DOMAINS.	
CC	-!- SIMILARITY: CONTAINS 2 LAMININ DOMAINS IV.	
CC	-!- SIMILARITY: CONTAINS 5 LAMININ G-LIKE DOMAINS.	
CC	-----	
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration	
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -	
CC	the European Bioinformatics Institute. There are no restrictions on its	
CC	use by non-profit institutions as long as its content is in no way	
CC	modified and this statement is not removed. Usage by and for commercial	
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/	
CC	or send an email to license@isb-sib.ch).	
CC	-----	
DR	EMBL; AJ293593; CAB99255.1; ..	
DR	EMBL; U37501; AAC53430.1; ..	
DR	HSP; P02468; LTLE.	
DR	MGI; MGI:105382; Lama5.	
DR	InterPro; IPR000561; EGF-like.	
DR	InterPro; IPR001886; LamNT.	
DR	InterPro; IPR000034; Laminin_B.	
DR	InterPro; IPR002049; Laminin_EGF.	
DR	InterPro; IPR001791; Laminin_G.	
DR	Pfam; PF00052; laminin_B; 1.	
DR	Pfam; PF00053; laminin_EGF; 19.	
DR	Pfam; PF00054; laminin_G; 2.	
DR	Pfam; PF00055; laminin_Nterm; 1.	
DR	ProDom; PD002082; LamNT; 1.	
DR	ProDom; PD003031; Laminin_B; 1.	
DR	PROSITE; PS00022; EGF_1; 19.	
DR	PROSITE; PS01186; EGF_2; 3.	
DR	PROSITE; PS01248; LAMININ_TYPE_EGF; 19.	
DR	PROSITE; PS00025; LAM_G_DOMAIN; 5.	
DR	Glycoprotein; Basement membrane; Extracellular matrix; Coiled coil;	
KW	Laminin EGF-like domain; Cell adhesion; Repeat; Signal.	
FT	SIGNAL	1 40
FT	CHAIN	41 3718
FT	DOMAIN	41 304
FT	DOMAIN	305 363
FT	DOMAIN	364 433
FT	DOMAIN	434 479
FT	DOMAIN	500 546
FT	DOMAIN	

```
FT DOMAIN 547 592 LAMININ EGF-LIKE 5. BY SIMILARITY. 1594
FT DOMAIN 593 637 LAMININ EGF-LIKE 6. BY SIMILARITY. 1582
FT DOMAIN 638 682 LAMININ EGF-LIKE 7. BY SIMILARITY. 1584
FT DOMAIN 683 728 LAMININ EGF-LIKE 8. BY SIMILARITY. 1601
FT DOMAIN 729 781 LAMININ EGF-LIKE 9. BY SIMILARITY. 1612
FT DOMAIN 782 833 LAMININ EGF-LIKE 10. BY SIMILARITY. 1630
FT DOMAIN 834 855 LAMININ EGF-LIKE 11 (INCOMPLETE). BY SIMILARITY. 1655
FT DOMAIN 856 1442 LAMININ DOMAIN IV 1 (DOMAIN IV B). BY SIMILARITY. 1881
FT DOMAIN 1443 1488 LAMININ EGF-LIKE 12. BY SIMILARITY. 1912
FT DOMAIN 1489 1532 LAMININ EGF-LIKE 13. BY SIMILARITY. 1930
FT DOMAIN 1533 1581 LAMININ EGF-LIKE 14. BY SIMILARITY. 1939
FT DOMAIN 1582 1632 LAMININ EGF-LIKE 15. BY SIMILARITY. 1950
FT DOMAIN 1633 1642 LAMININ EGF-LIKE 16 (N-TERMINAL). BY SIMILARITY. 1968
FT DOMAIN 1643 1831 LAMININ DOMAIN IV 2 (DOMAIN IV A). BY SIMILARITY. 1971
FT DOMAIN 1832 1864 LAMININ EGF-LIKE 17 (C-TERMINAL). BY SIMILARITY. 1986
FT DOMAIN 1865 1914 LAMININ EGF-LIKE 18. BY SIMILARITY. 1993
FT DOMAIN 1915 1970 LAMININ EGF-LIKE 19. BY SIMILARITY. 2005
FT DOMAIN 1971 2024 LAMININ EGF-LIKE 20. BY SIMILARITY. 2022
FT DOMAIN 2025 2071 LAMININ EGF-LIKE 21. BY SIMILARITY. 2083
FT DOMAIN 2072 2118 LAMININ EGF-LIKE 22. BY SIMILARITY. 2090
FT DOMAIN 2119 2168 LAMININ EGF-LIKE 23. BY SIMILARITY. 2101
FT DOMAIN 2169 2735 LAMININ II AND I. BY SIMILARITY. 2116
FT DOMAIN 2736 2933 LAMININ G-LIKE 1. BY SIMILARITY. 2126
FT DOMAIN 2947 3119 LAMININ G-LIKE 2. BY SIMILARITY. 2133
FT DOMAIN 3128 3296 LAMININ G-LIKE 3. BY SIMILARITY. 2144
FT DOMAIN 3337 3511 LAMININ G-LIKE 4. BY SIMILARITY. 2166
FT DOMAIN 3518 3689 LAMININ G-LIKE 5. BY SIMILARITY. 2172
FT DOMAIN 2205 2257 COILED COIL (POTENTIAL). BY SIMILARITY. 2172
FT DOMAIN 2258 2464 COILED COIL (POTENTIAL). BY SIMILARITY. 2172
FT DOMAIN 2465 2604 COILED COIL (POTENTIAL). BY SIMILARITY. 2172
FT DOMAIN 2605 2639 CELL ATTACHMENT SITE (POTENTIAL). BY SIMILARITY. 2172
FT SITE 1723 1725 CELL ATTACHMENT SITE (POTENTIAL). BY SIMILARITY. 2172
FT SITE 1839 1841 BY SIMILARITY. 2172
FT DISULFID 305 314 BY SIMILARITY. 2172
FT DISULFID 307 327 BY SIMILARITY. 2172
FT DISULFID 329 338 BY SIMILARITY. 2172
FT DISULFID 341 361 BY SIMILARITY. 2172
FT DISULFID 364 373 BY SIMILARITY. 2172
FT DISULFID 366 398 BY SIMILARITY. 2172
FT DISULFID 401 410 BY SIMILARITY. 2172
FT DISULFID 413 431 BY SIMILARITY. 2172
FT DISULFID 434 445 BY SIMILARITY. 2172
FT DISULFID 436 452 BY SIMILARITY. 2172
FT DISULFID 454 463 BY SIMILARITY. 2172
FT DISULFID 466 476 BY SIMILARITY. 2172
FT DISULFID 500 512 BY SIMILARITY. 2172
FT DISULFID 523 532 BY SIMILARITY. 2172
FT DISULFID 535 544 BY SIMILARITY. 2172
FT DISULFID 547 559 BY SIMILARITY. 2172
FT DISULFID 568 577 BY SIMILARITY. 2172
FT DISULFID 580 590 BY SIMILARITY. 2172
FT DISULFID 593 605 BY SIMILARITY. 2172
FT DISULFID 595 611 BY SIMILARITY. 2172
FT DISULFID 613 622 BY SIMILARITY. 2172
FT DISULFID 625 635 BY SIMILARITY. 2172
FT DISULFID 638 650 BY SIMILARITY. 2172
FT DISULFID 640 656 BY SIMILARITY. 2172
FT DISULFID 658 667 BY SIMILARITY. 2172
FT DISULFID 670 680 BY SIMILARITY. 2172
FT DISULFID 683 695 BY SIMILARITY. 2172
FT DISULFID 695 702 BY SIMILARITY. 2172
FT DISULFID 704 713 BY SIMILARITY. 2172
FT DISULFID 716 726 BY SIMILARITY. 2172
FT DISULFID 1443 1455 BY SIMILARITY. 2172
FT DISULFID 1445 1462 BY SIMILARITY. 2172
FT DISULFID 1464 1473 BY SIMILARITY. 2172
FT DISULFID 1476 1486 BY SIMILARITY. 2172
FT DISULFID 1533 1548 BY SIMILARITY. 2172
FT DISULFID 1555 1566 BY SIMILARITY. 2172
FT DISULFID 1569 1579 BY SIMILARITY. 2172

Query Match 8.6%; Score 202.5; DB 1; Length 3718;
Best Local Similarity 20.0%; Pred. No. 3.4e-06;
Matches 94; Conservative 36; Mismatches 150; Indels 191; Gaps 21;

QY 43 TCVNCRPNY-----YNGGAAQGEANGNQPPFAANNAARGICVPCQINRVGSVTN----- 91
Db 1842 SCQECAPGYRDTKGLFGRVPCQCHGHSRCLPGS--GICVGCQHTGDCQECRCRP 1899

QY 92 --AGDLATATOC-STQCP-----TGATLDGVTDFDR-----SAAQCVKCKPNFY 136
Db 1900 FVSSDFSNFASPCVPCPLAVPSNFDGCVLRNGRTQCLCRPGYAGASCERCAPGFF- 1958

QY 137 NGSPQGEAPGVQVFAAGAAAAGVAATVSCVPCOLNKNDSPATAGAAANLATOC SNQCP 196
Db 1959 --GNP-----LVLGSSQPCDCSNGSDPNMIFSDCDPLTGACRGC- 1996

QY 197 TGTVLDDGVTLYVFNISALCVKCRPNFYNGSGPQGEAPGVQVFAAGAAAAGVAATVSC 256
Db 1997 -----LRHTTGPHCERCAPGY-----GNALLPGNC 2022

QY 257 VPCQINKNDSPATAGAAANLATOCSTQCTGTATOD-----GVTLVFSNHSSTQCS 306
Db 2023 TRC-----DCS---PCGTETCDPQSGRCLKAGVT-----GORCD 2054

QY 307 QCIAANYFNGNFEACKSQCLKC-----PVSKTTPAH-----APGNTATQATQCL-- 350
Db 2055 RCLGYF-----GFECQCGCRPCACGPAKSGECPHQSQCHQCPGTTGPOCLECAPG 2107

QY 351 -----TTCPAGTVLDDGTSNFINVASATECTKCSAGF---F 382
Db 2108 YWGLPEKCRQCQPRGHCDPHTGHCCTPPG-----LSGERCDTCSQQHQVY 2155

QY 383 ASKTTGTAGTDTCTECTKKL-----TSGATAKVAEATOKVOCASSTFEAK 428
```

Db 2156 PKPGRGHICEVCHCIVLLDRLERAGALLPAIREQLGINASSAAWAR 2206

RESULT 3

```
LMAL_MOUSE
ID LMAL_MOUSE STANDARD; PRT; 3084 AA.
AC P19137;
DT 01-NOV-1990 (Rel. 16, Created)
DT 01-NOV-1990 (Rel. 16, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Laminin alpha-1 chain precursor (Laminin A chain).
GN LMAL1 OR LAMA-1 OR LAMA.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RX MEDLINE=89034134; PubMed=3182802;
RA Sasaki M., Kleinman H.K., Huber H., Deutzmann R., Yamada Y.;
RT "Laminin, a multidomain protein. The A chain has a unique globular
RT domain and homology with the basement membrane proteoglycan and the
RT laminin B chains.";
RL J. Biol. Chem. 263:16536-16544(1988).
RN [2]
RX MEDLINE=89034134; PubMed=3182802;
RA Hartl L., Oberbaumer I., Deutzmann R.;
RT "The N terminus of laminin A chain is homologous to the B chains.";
RL Eur. J. Biochem. 173:629-635(1988).
RN [3]
RX MEDLINE=89030693; PubMed=3181157;
RA Deutzmann R., Huber J., Schmetz K.A., Oberbaumer I., Hartl L.;
RT "Structural study of long arm fragments of laminin. Evidence for
RT repetitive C-terminal sequences in the A-chain, not present in the B-
RT chains.";
RL Eur. J. Biochem. 177:35-45(1988).
CC -!- FUNCTION: Binding to cells via a high affinity receptor, laminin
CC is thought to mediate the attachment, migration, and organization
CC of cells into tissues during embryonic development by interacting
CC with other extracellular matrix components.
CC -!- SUBUNIT: Laminin is a complex glycoprotein, consisting of three
CC different polypeptide chains (alpha, beta, gamma), which are bound
CC to each other by disulfide bonds into a cross-shaped molecule
CC Comprising one long and three short arms with globules at each
CC end.
CC THE ALPHA-1 CHAIN IS A SUBUNIT OF LAMININ-1 (EHS LAMININ) AND
CC LAMININ-3 (S-LAMININ).
CC -!- SUBCELLULAR LOCATION: Extracellular.
CC -!- TISSUE SPECIFICITY: FOUND IN THE BASEMENT MEMBRANES (MAJOR
CC COMPONENT).
CC -!- DOMAIN: THE ALPHA-HELICAL DOMAINS I AND II ARE THOUGHT TO INTERACT
CC WITH OTHER LAMININ CHAINS TO FORM A COILED COIL STRUCTURE.
CC -!- DOMAIN: DOMAINS VI, IV AND G ARE GLOBULAR.
CC -!- SIMILARITY: CONTAINS 1 LAMININ N-TERMINAL DOMAIN (DOMAIN VI).
CC -!- SIMILARITY: CONTAINS 17 LAMININ EGF-LIKE DOMAINS.
CC -!- SIMILARITY: CONTAINS 2 LAMININ DOMAINS IV.
CC -!- SIMILARITY: CONTAINS 5 LAMININ G-LIKE DOMAINS.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL; J04064; AAA39410.1; -
DR EMBL; X07737; CAA30561.1; -
DR EMBL; X13459; CAA31807.1; -
DR EMBL; M36775; AAA39406.1; -
DR PIR; A31771; MMSA.
```

```
HSSP: Q60675; IQUO.
DR MGD; MGI:99892; Lama1.
DR InterPro; IPR000561; EGF-like.
DR InterPro; IPR001886; LamNT.
DR InterPro; IPR000034; Laminin_B.
DR InterPro; IPR002049; Laminin_EGF.
DR InterPro; IPR001791; Laminin_G.
DR Pfam; PF00052; laminin_B; 2.
DR Pfam; PF00053; laminin_EGF; 15.
DR Pfam; PF00054; laminin_G; 5.
DR Pfam; PF00055; laminin_Nterm; 1.
DR PRINTS; PR00011; EGF_LAMININ.
DR PRODOM; PD002082; LamNT; 1.
DR PRODOM; PD003031; Laminin_B; 2.
DR SMART; SM00180; EGF_Lam; 14.
DR SMART; SM00281; Lamb; 2.
DR SMART; SM00282; LamG; 5.
DR SMART; SM00382; LamNT; 1.
DR PROSITE; PS00022; EGF_1; 11.
DR PROSITE; PS01186; EGF_2; 3.
DR PROSITE; PS01248; LAMININ_TYPE_EGF; 15.
DR PROSITE; PS50025; LAM_G_DOMAIN; 5.
DR Glycoprotein; Basement membrane; Extracellular matrix; Coiled coil;
KW Laminin EGF-like domain; Cell adhesion; Repeat; Signal.
FT SIGNAL 1 24 POTENTIAL.
FT CHAIN 25 3084 LAMININ ALPHA-1 CHAIN.
FT MOD_RES 25 25 BLOCKED.
FT DOMAIN 25 276 LAMININ N-TERMINAL (DOMAIN VI).
FT DOMAIN 277 333 LAMININ EGF-LIKE 1.
FT DOMAIN 334 403 LAMININ EGF-LIKE 2.
FT DOMAIN 404 460 LAMININ EGF-LIKE 3.
FT DOMAIN 461 509 LAMININ EGF-LIKE 4.
FT DOMAIN 510 519 LAMININ EGF-LIKE 5 (N-TERMINAL).
FT DOMAIN 520 715 LAMININ DOMAIN IV 1 (DOMAIN IV B).
FT DOMAIN 716 748 LAMININ EGF-LIKE 5 (C-TERMINAL).
FT DOMAIN 749 797 LAMININ EGF-LIKE 6.
FT DOMAIN 798 855 LAMININ EGF-LIKE 7.
FT DOMAIN 856 908 LAMININ EGF-LIKE 8.
FT DOMAIN 909 957 LAMININ EGF-LIKE 9.
FT DOMAIN 958 1004 LAMININ EGF-LIKE 10.
FT DOMAIN 1005 1050 LAMININ EGF-LIKE 11.
FT DOMAIN 1051 1096 LAMININ EGF-LIKE 12.
FT DOMAIN 1097 1156 LAMININ EGF-LIKE 13.
FT DOMAIN 1157 1166 LAMININ EGF-LIKE 14 (N-TERMINAL).
FT DOMAIN 1167 1368 LAMININ DOMAIN IV 2 (DOMAIN IV A).
FT DOMAIN 1369 1409 LAMININ EGF-LIKE 14 (C-TERMINAL).
FT DOMAIN 1410 1458 LAMININ EGF-LIKE 15.
FT DOMAIN 1459 1515 LAMININ EGF-LIKE 16.
FT DOMAIN 1516 1562 LAMININ EGF-LIKE 17.
FT DOMAIN 1563 2124 DOMAIN II AND I.
FT DOMAIN 2125 2305 LAMININ G-LIKE 1.
FT DOMAIN 2306 2313 LAMININ G-LIKE 2.
FT DOMAIN 2314 2889 LAMININ G-LIKE 3.
FT DOMAIN 2890 2894 LAMININ G-LIKE 4.
FT DOMAIN 2895 2899 LAMININ G-LIKE 5.
FT DOMAIN 2900 3079 COILED COIL (POTENTIAL).
FT DOMAIN 3080 3086 COILED COIL (POTENTIAL).
FT DOMAIN 3087 3093 COILED COIL (POTENTIAL).
FT DOMAIN 3094 3100 COILED COIL (POTENTIAL).
FT SITE 1147 1149 CELL ATTACHMENT SITE.
FT DISULFID 277 286 BY SIMILARITY.
FT DISULFID 279 297 BY SIMILARITY.
FT DISULFID 299 308 BY SIMILARITY.
FT DISULFID 311 331 BY SIMILARITY.
FT DISULFID 334 343 BY SIMILARITY.
FT DISULFID 336 368 BY SIMILARITY.
FT DISULFID 371 380 BY SIMILARITY.
FT DISULFID 383 401 BY SIMILARITY.
FT DISULFID 404 416 BY SIMILARITY.
FT DISULFID 406 434 BY SIMILARITY.
FT DISULFID 436 445 BY SIMILARITY.
FT DISULFID 448 458 BY SIMILARITY.
FT DISULFID 461 474 BY SIMILARITY.
FT DISULFID 463 478 BY SIMILARITY.
```

FT DISULFID 480 489 BY SIMILARITY.
 FT DISULFID 492 507 BY SIMILARITY.
 FT DISULFID 749 758 BY SIMILARITY.
 FT DISULFID 751 764 BY SIMILARITY.
 FT DISULFID 767 776 BY SIMILARITY.
 FT DISULFID 779 795 BY SIMILARITY.
 FT DISULFID 798 813 BY SIMILARITY.
 FT DISULFID 800 823 BY SIMILARITY.
 FT DISULFID 826 835 BY SIMILARITY.
 FT DISULFID 838 853 BY SIMILARITY.
 FT DISULFID 856 870 BY SIMILARITY.
 FT DISULFID 858 877 BY SIMILARITY.
 FT DISULFID 880 889 BY SIMILARITY.
 FT DISULFID 892 906 BY SIMILARITY.
 FT DISULFID 909 921 BY SIMILARITY.
 FT DISULFID 911 928 BY SIMILARITY.
 FT DISULFID 930 939 BY SIMILARITY.
 FT DISULFID 942 955 BY SIMILARITY.
 FT DISULFID 958 970 BY SIMILARITY.
 FT DISULFID 960 976 BY SIMILARITY.
 FT DISULFID 978 987 BY SIMILARITY.
 FT DISULFID 990 1002 BY SIMILARITY.
 FT DISULFID 1005 1014 BY SIMILARITY.
 FT DISULFID 1007 1021 BY SIMILARITY.
 FT DISULFID 1023 1032 BY SIMILARITY.
 FT DISULFID 1035 1048 BY SIMILARITY.
 FT DISULFID 1051 1063 BY SIMILARITY.
 FT DISULFID 1053 1070 BY SIMILARITY.
 FT DISULFID 1072 1081 BY SIMILARITY.
 FT DISULFID 1084 1094 BY SIMILARITY.
 FT DISULFID 1410 1419 BY SIMILARITY.
 FT DISULFID 1412 1426 BY SIMILARITY.
 FT DISULFID 1429 1438 BY SIMILARITY.
 FT DISULFID 1441 1456 BY SIMILARITY.
 FT DISULFID 1449 1473 BY SIMILARITY.
 FT DISULFID 1461 1483 BY SIMILARITY.
 FT DISULFID 1485 1495 BY SIMILARITY.
 FT DISULFID 1498 1513 BY SIMILARITY.
 FT DISULFID 1516 1528 BY SIMILARITY.
 FT DISULFID 1518 1535 BY SIMILARITY.
 FT DISULFID 1537 1546 BY SIMILARITY.
 FT DISULFID 1549 1560 BY SIMILARITY.
 FT DISULFID 1563 1563 INTERCHAIN (PROBABLE).
 FT DISULFID 1567 1567 INTERCHAIN (PROBABLE).
 FT CARBOHYD 43 45 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 79 79 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 370 370 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 374 374 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 531 531 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 562 562 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 672 672 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 770 770 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 857 857 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 914 914 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 959 959 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 969 969 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1052 1052 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1344 1344 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1414 1414 N-LINKED (GLCNAC. . .) (POTENTIAL).
 Query Match 8.5%; Score 198; DB 1; Length 3084;
 Best Local Similarity 20.0%; Pred. No. 5.8e-06;
 Matches 89; Conservative 38; Mismatches 142; Indels 176; Gaps 16;
 QY 23 CPDG-TQTQAGLTDVGAADLGTVCNCRPNFYNGGAQGEANGNPFAANNAARGICVPC 81
 DB 718 CQGGTGT-----SCACLPGYRVDGILFG-----GICQPC 749
 QY 82 QINRVGSVTNAGDLATLATOCST-----QCPTG----- 109
 DB 750 ECHGHASECDTHGICSVCTHTTGDHCEQCLPGFYGTFPSRGTGDCQPCACPLSIDSNFN 809
 QY 110 ---TALDGVTDVDFR-----SAAQCVKCKPNFYNGSGSPQGEAPGVQVFAAGAAAGVA 161

Db 810 SPTCHLTDGEEVVDQCAPGYSGSMCERCADGYGNTPVGG----- 851
 QY 162 AVTSQVPCQLKNDKNDSPATAGAQAANLATQCSNOCPTGTVLDDGVTLVFNITSATLCVKCRP 221
 Db 852 ----TCVPCNCSGNVDPLEAGHCDSDVTGEC-----LKCLWNTDGAHCERCAD 894
 QY 222 NFYYNGSGSQGEAPGVQVFAAGAAAGVAAVTSQVPCQLKNDKNDSPATAGAQAANLATQCS 281
 Db 895 GFY-----GDVATKNCRCADCHENG-----LS 918
 QY 282 TQCPGTGTAIOGVTVLFSNSTQCSQCIANYFFNGFNEACKSQCLKCPVSKTTPAHAPGN 341
 Db 919 GVCHLETGLCDCKPHV---TGQCDQCLSGY-----GLDTGLGCVPCNCSVEGVSVD 968
 QY 342 TATQATOCCLTCTCPAGTVLDDGTSTNFNVASATECTKCSAGFFASKTTGTFAGTDTCTECHK 401
 Db 969 NCTEEOGC--HCGPG-----VSGKQCDRCSHGFYAFQDGG-----CTPCDC 1007
 QY 402 KLTSGATAKYAEA-----TQKVQC 421
 Db 1008 AHTQNNCDPASGEGCLCPHTQGLKC 1032
 RESULT 4
 LMA2_HUMAN
 ID LMA2_HUMAN STANDARD; PRT: 3110 AA.
 AC P24043; Q14736; Q93022;
 DT 01-MAR-1992 (Rel. 21, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Laminin alpha-2 chain precursor (Laminin M chain) (Merosin heavy chain).
 DE chain).
 GN LAMA2 OR LAMM.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_Taxid=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Placenta;
 RX MEDLINE=94124633; PubMed=8294519;
 RA Vuolteenaho R., Nissinen M., Sainio K., Byers M., Eddy R.,
 RA Hirvonen H., Shows T.B., Sariola H., Engvall E., Tryggvason K.;
 RT "Human laminin M chain (merosin): complete primary structure,
 RT chromosomal assignment, and expression of the M and A chain in human
 RT fetal tissues."; J. Cell Biol. 124:381-394(1994).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=97066955; PubMed=8910357;
 RA Zhang X., Vuolteenaho R., Tryggvason K.;
 RT "Structure of the human laminin alpha2-chain gene (LAMA2), which is
 RT affected in congenital muscular dystrophy."; J. Biol. Chem. 271:27664-27669(1996).
 RN [3]
 RP SEQUENCE OF 1981-3110 FROM N.A., AND PARTIAL SEQUENCE.
 RC TISSUE=Placenta;
 RX MEDLINE=90238994; PubMed=2185464;
 RA Ehrig K., Leivo I., Argraves W.S., Ruoslahti E., Engvall E.;
 RT "Merosin, a tissue-specific basement membrane protein, is a
 RT laminin-like protein."; Proc. Natl. Acad. Sci. U.S.A. 87:3264-3268(1990).
 RN [4]
 RP VARIANTS GLN-545; HIS-619; LEU-919; HIS-2586 AND LYS-2614.
 RA Panicker S.G., Mendell J.T., Chen L., Feng B., Sahenk Z.,
 RA Marzluf G.A., Anato A.A., Mendell J.R.;
 RT "Novel single base polymorphisms and rare sequence variants in
 RT the laminin 2-chain coding region detected by RNA/SSCP analysis."; Hum. Mutat. 13:174-174(1999).
 RN [5]
 RP ERRATUM.
 RA Panicker S.G., Mendell J.T., Chen L., Feng B., Sahenk Z.,

RA Marzluf G.A., Amato A.A., Mendell J.R.;
RL Hum. Mutat. 13:340-340(1999).
CC -|- FUNCTION: Binding to cells via a high affinity receptor, laminin
CC is thought to mediate the attachment, migration, and organization
CC of cells into tissues during embryonic development by interacting
CC with other extracellular matrix components.
CC -|- SUBUNIT: Laminin is a complex glycoprotein, consisting of three
CC different polypeptide chains (alpha, beta, gamma), which are bound
CC to each other by disulfide bonds into a cross-shaped molecule
CC comprising one long and three short arms with globules at each
CC end.
CC THE ALPHA-2 CHAIN IS A SUBUNIT OF LAMININ-2 (MEROSIN) AND LAMININ-
CC 4 (S-MEROSIN).
CC -|- SUBCELLULAR LOCATION: EXTRACELLULAR; FOUND IN THE BASEMENT
CC MEMBRANES (MAJOR COMPONENT).
CC -|- TISSUE SPECIFICITY: PLACENTA, STRIATED MUSCLE, PERIPHERAL NERVE,
CC CARDIAC MUSCLE, PANCREAS, LUNG, SPLEEN, KIDNEY, ADRENAL GLAND,
CC SKIN, TESTIS, MENINGES, CHOROID PLEXUS, AND SOME OTHER REGIONS OF
CC THE BRAIN; NOT IN LIVER, THYMUS AND BONE.
CC -|- DOMAIN: THE ALPHA-HELICAL DOMAINS I AND II ARE THOUGHT TO INTERACT
CC WITH OTHER LAMININ CHAINS TO FORM A COILED COIL STRUCTURE.
CC -|- DOMAIN: DOMAINS VI, IV AND G ARE GLOBULAR.
CC -|- DISEASE: DEFECTS IN LAMA2 ARE THE CAUSE OF MEROSIN-DEFICIENT
CC CONGENITAL MUSCULAR DYSTROPHY (MCMD).
CC -|- SIMILARITY: CONTAINS 17 LAMININ N-TERMINAL DOMAIN (DOMAIN VI).
CC -|- SIMILARITY: CONTAINS 2 LAMININ EGF-LIKE DOMAINS.
CC -|- SIMILARITY: CONTAINS 5 LAMININ G-LIKE DOMAINS.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>)
CC or send an email to license@isb-sib.ch).

DR EMBL; 22653; CRA81394.1; .
DR EMBL; U66796; AAB18388.1; .
DR EMBL; U66733; AAB18388.1; JOINED.
DR EMBL; U66734; AAB18388.1; JOINED.
DR EMBL; U66735; AAB18388.1; JOINED.
DR EMBL; U66736; AAB18388.1; JOINED.
DR EMBL; U66737; AAB18388.1; JOINED.
DR EMBL; U66738; AAB18388.1; JOINED.
DR EMBL; U66739; AAB18388.1; JOINED.
DR EMBL; U66740; AAB18388.1; JOINED.
DR EMBL; U66741; AAB18388.1; JOINED.
DR EMBL; U66742; AAB18388.1; JOINED.
DR EMBL; U66743; AAB18388.1; JOINED.
DR EMBL; U66745; AAB18388.1; JOINED.
DR EMBL; U66746; AAB18388.1; JOINED.
DR EMBL; U66747; AAB18388.1; JOINED.
DR EMBL; U66748; AAB18388.1; JOINED.
DR EMBL; U66749; AAB18388.1; JOINED.
DR EMBL; U66750; AAB18388.1; JOINED.
DR EMBL; U66751; AAB18388.1; JOINED.
DR EMBL; U66752; AAB18388.1; JOINED.
DR EMBL; U66753; AAB18388.1; JOINED.
DR EMBL; U66754; AAB18388.1; JOINED.
DR EMBL; U66755; AAB18388.1; JOINED.
DR EMBL; U66756; AAB18388.1; JOINED.
DR EMBL; U66757; AAB18388.1; JOINED.
DR EMBL; U66758; AAB18388.1; JOINED.
DR EMBL; U66759; AAB18388.1; JOINED.
DR EMBL; U66760; AAB18388.1; JOINED.
DR EMBL; U66761; AAB18388.1; JOINED.
DR EMBL; U66762; AAB18388.1; JOINED.
DR EMBL; U66763; AAB18388.1; JOINED.
DR EMBL; U66764; AAB18388.1; JOINED.
DR EMBL; U66765; AAB18388.1; JOINED.
DR EMBL; U66766; AAB18388.1; JOINED.
DR EMBL; U66768; AAB18388.1; JOINED.
EMBL; U66769; AAB18388.1; JOINED.
EMBL; U66770; AAB18388.1; JOINED.
EMBL; U66771; AAB18388.1; JOINED.
EMBL; U66772; AAB18388.1; JOINED.
EMBL; U66773; AAB18388.1; JOINED.
EMBL; U66774; AAB18388.1; JOINED.
EMBL; U66775; AAB18388.1; JOINED.
EMBL; U66776; AAB18388.1; JOINED.
EMBL; U66777; AAB18388.1; JOINED.
EMBL; U66778; AAB18388.1; JOINED.
EMBL; U66779; AAB18388.1; JOINED.
EMBL; U66780; AAB18388.1; JOINED.
EMBL; U66781; AAB18388.1; JOINED.
EMBL; U66782; AAB18388.1; JOINED.
EMBL; U66783; AAB18388.1; JOINED.
EMBL; U66784; AAB18388.1; JOINED.
EMBL; U66785; AAB18388.1; JOINED.
EMBL; U66786; AAB18388.1; JOINED.
EMBL; U66787; AAB18388.1; JOINED.
EMBL; U66788; AAB18388.1; JOINED.
EMBL; U66789; AAB18388.1; JOINED.
EMBL; U66790; AAB18388.1; JOINED.
EMBL; U66791; AAB18388.1; JOINED.
EMBL; U66792; AAB18388.1; JOINED.
EMBL; U66793; AAB18388.1; JOINED.
EMBL; U66794; AAB18388.1; JOINED.
EMBL; U66795; AAB18388.1; JOINED.
EMBL; M59832; AAA63215.1; .
PIR; A35899; MMHUMH.
HSP; Q60675; IQUO.
Gene; HGNC:6482; LAMA2.
MIM; 156225; .
InterPro; IPR000561; EGF-like.
InterPro; IPR001886; LamNT.
InterPro; IPR000034; Laminin_B.
InterPro; IPR002049; Laminin_EGF.
InterPro; IPR001791; Laminin_G.
Pfam; PF00052; laminin_B; 2.
Pfam; PF00053; laminin_EGF; 15.
Pfam; PF00054; laminin_G; 5.
Pfam; PF00055; laminin_Nterm; 1.
PRINTS; PR00011; EGF_LAMININ.
ProDom; PD002082; LamNT; 1.
SMART; SM00180; EGF_Lam; 14.
SMART; SM00281; Lamb; 2.
SMART; SM00282; Lamb; 5.
SMART; SM00136; LamNT; 1.
PROSITE; PS00022; EGF_1; 11.
PROSITE; PS01186; EGF_2; 3.
PROSITE; PS01248; LAMININ_TYPE_EGF; 14.
PROSITE; PS00025; LAM_G_DOMAIN; 5.
KW Glycoprotein; Basement membrane; Extracellular matrix; Coiled coil;
KW Laminin EGF-like domain; Cell adhesion; Repeat; Signal; Polymorphism.
FT SIGNAL 1 22 POTENTIAL.
FT CHAIN 23 3110 LAMININ ALPHA-2 CHAIN.
FT DOMAIN 23 286 LAMININ N-TERMINAL (DOMAIN VI).
FT DOMAIN 287 343 LAMININ EGF-LIKE 1.
FT DOMAIN 344 413 LAMININ EGF-LIKE 2.
FT DOMAIN 414 468 LAMININ EGF-LIKE 3.
FT DOMAIN 469 517 LAMININ EGF-LIKE 4.
FT DOMAIN 518 527 LAMININ EGF-LIKE 5 (N-TERMINAL).
FT DOMAIN 528 723 LAMININ DOMAIN IV 1 (DOMAIN IV B).
FT DOMAIN 724 756 LAMININ EGF-LIKE 5 (C-TERMINAL).
FT DOMAIN 757 806 LAMININ EGF-LIKE 6.
FT DOMAIN 807 864 LAMININ EGF-LIKE 7.
FT DOMAIN 865 917 LAMININ EGF-LIKE 8.
FT DOMAIN 918 966 LAMININ EGF-LIKE 9.
FT DOMAIN 967 1013 LAMININ EGF-LIKE 10.
FT DOMAIN 1014 1059 LAMININ EGF-LIKE 11.
FT DOMAIN 1060 1105 LAMININ EGF-LIKE 12.
FT DOMAIN 1106 1165 LAMININ EGF-LIKE 13.

```

FT DOMAIN 1166 1175 LAMININ EGF-LIKE 14 (N-TERMINAL).
FT DOMAIN 1176 1379 LAMININ DOMAIN IV 2 (DOMAIN IV A).
FT DOMAIN 1380 1419 LAMININ EGF-LIKE 14 (C-TERMINAL).
FT DOMAIN 1420 1468 LAMININ EGF-LIKE 15.
FT DOMAIN 1469 1526 LAMININ EGF-LIKE 16.
FT DOMAIN 1527 1573 LAMININ EGF-LIKE 17.
FT DOMAIN 1574 2144 DOMAIN II AND I.

Query Match 8.2%; Score 191; DB 1; Length 3110;
Best Local Similarity 24.4%; Pred. No. 1.8e-05;
Matches 113; Conservative 34; Mismatches 158; Indels 158; Gaps 30;

QY 42 GTCVNCR-----PNFYNGGAAGGANGNQFAA--NNAARICVPCQINRVG 87
Db 771 GECLNCKDHTGGPYCKLPGFY--CEPTKGTSEDQPCACPLNIPSNFSTCHLDR-- 826
QY 88 SVTNAGDLATLQCSTQCTPTGTALDDGVTVDFDRSAAQCVKCKPNFYNGSGPOGEAP 147
Db 827 -----SLGLICD-GCPVG-----YTGPRCERCAEGYF--GQP--SVPG 859
QY 148 VOVFAAGAAAGVAAVTSOCVPCOLNKNKNDSPATAGAAQANLATQCSNQCPTGT 200
Db 860 -----GSCQPCQCNNDLDFIPGSCDSLGSCL-ICKPGTTGRYCELC 901
QY 201 -----LDGVTLVNTSATLCVCRPNFYNGG-----SPOGEA--PGVQVFAAGAA 246
Db 902 ADGYFGDAV-----DAKNCQPCRCN---AGGSFSEVCHSQTGCCECRANVQGRCKCK 952
QY 247 AGVAAVTSQ--CVPCQINKNDSPATAGAAQANLATQCSQCTPTGTATQGVTVFSN 304
Db 953 AGTFGLQARGVPCNCFNFGS-----KSPDCESGQC-----WCQPGVT-----GKK 995
QY 305 CSCCIANYFNGNFAGKSCQLKCPVS-----KTPPAHAPGNT-ATQATQC----- 349
Db 996 CDRCAHYF---NFQEG--GCTACECSHLGNNDCKPCKRCICPPNTIGERCKCAPNTWG 1050
QY 350 ---LTT---CPAGTV---LDDGTSTNF-----VASATECTKCSAGFF-----ASKTTGFT 390
Db 1051 HSITGCKKACNSCTVGSDFQCNVNTGQCNCHPKFSKAKTCSRGHWNYPKCNLCDCFL 1110
QY 391 AGTD--TCTECTKKL-----TSGATAKYAEATQKVCASITF 426
Db 1111 PGTDATTCDSERKKKSCSDQTGCTCKVNVVGEIHCDRCRPOKF 1153

RESULT 5
ZAN_MOUSE STANDARD; PRT; 5376 AA.
AC O88799; O08647;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DE Zonadhesin precursor.
GN ZAN.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Testis;
RX MEDLINE=98123114; PubMed=9452463;
RA Gao Z., Garbers D.L.;
RT "Species diversity in the structure of zonadhesin, a sperm-specific
RT membrane protein containing multiple cell adhesion molecule-like
RL domains.";
RL J. Biol. Chem. 273:3415-3421(1998).
RN [2]
RP SEQUENCE OF 4864-5376 FROM N.A.
RC TISSUE=Testis;
RX MEDLINE=97271566; PubMed=9126492;
RA Gao Z., Harumi T., Garbers D.L.;
RT "Chromosome localization of the mouse zonadhesin gene and the human

```

```

zonadhesin gene (ZAN).";
Genomics 41:119-122(1997).
-1- FUNCTION: BINDS IN A SPECIES-SPECIFIC MANNER TO THE ZONA PELLUCIDA
OF THE EGG. MAY BE INVOLVED IN GAMETE RECOGNITION AND/OR
SIGNALING.
-1- SUBUNIT: PROBABLY FORMS COVALENT OLIGOMERS.
-1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN, EXCLUSIVELY ON THE
APICAL REGION OF THE SPERM HEAD.
-1- TISSUE SPECIFICITY: IN TESTIS; PRIMARILY IN HAPLOID SPERMATIDS.
-1- DOMAIN: THE MAM DOMAINS PROBABLY MEDIATES SPERM ADHESION TO THE
ZONA PELLUCIDA.
-1- DOMAIN: DURING SPERM MIGRATION THROUGH THE REPRODUCTIVE TRACTS,
THE MUCIN-LIKE DOMAIN MIGHT INHIBIT INAPPROPRIATE TRAPPING OF
SPERMATOZOEA OR PROMOTING ADHESION TO THE OVIDUCTAL ISTHMUS.
-1- DOMAIN: THE VWFD DOMAIN 2 MAY MEDIATE COVALENT
OLIGOMERIZATION (BY SIMILARITY TO HUMAN INTESTINAL MUCIN MUC2).
-1- SIMILARITY: CONTAINS 3 MAM DOMAINS.
-1- SIMILARITY: CONTAINS 25 VWFD DOMAINS.
-1- SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch)
-----
EMBL; U97068; AAC26680.1; -
DR EMBL; U83190; AAC53125.1; -
DR MGD; MGI:106656; ZAN.
DR InterPro; IPR000561; EGF-like.
DR InterPro; IPR003645; FOLN.
DR InterPro; IPR000998; MAM domain.
DR InterPro; IPR002919; TIL_Cysrich.
DR InterPro; IPR003328; TILA_Cysrich.
DR InterPro; IPR001007; VWF_C.
DR InterPro; IPR001846; VWF_D.
DR Pfam; PF00094; vwd; 4.
DR Pfam; PF00629; MAM; 3.
DR Pfam; PF01826; TIL; 25.
DR Pfam; PF02345; TILA; 25.
DR SMART; SM00181; EGF; 2.
DR SMART; SM00274; FOLN; 11.
DR SMART; SM00137; MAM; 2.
DR SMART; SM00214; VWC; 17.
DR SMART; SM00216; VWD; 4.
DR PROSITE; PS00022; EGF_1; 1.
DR PROSITE; PS01186; EGF_2; 18.
DR PROSITE; PS00740; MAM_1; FALSE_NEG.
DR PROSITE; PS50060; MAM_2; 3.
KW Signal; Glycoprotein; Transmembrane; EGF-like domain; Cell adhesion;
Repeat.
FT SIGNAL 1 17 POTENTIAL.
FT CHAIN 18 5376 ZONADHESIN.
FT DOMAIN 18 5310 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 5311 5337 POTENTIAL.
FT DOMAIN 5338 5376 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 45 210 MAM 1.
FT DOMAIN 215 374 MAM 2.
FT DOMAIN 377 542 MAM 3.
FT DOMAIN 547 1170 80 X HEPTAPEPTIDE REPEATS (APPROXIMATE)
(MUCIN-LIKE DOMAIN).
FT DOMAIN 1171 1280 VWFD 1 (PARTIAL).
FT DOMAIN 1281 1669 VWFD 2.
FT DOMAIN 1670 2056 VWFD 3.
FT DOMAIN 2057 2459 VWFD 4.
FT DOMAIN 2460 2579 VWFD 5 (PARTIAL).
FT DOMAIN 2580 2699 VWFD 6 (PARTIAL).
FT DOMAIN 2700 2819 VWFD 7 (PARTIAL).
FT DOMAIN 2820 2939 VWFD 8 (PARTIAL).
FT DOMAIN 2940 3059 VWFD 9 (PARTIAL).
FT DOMAIN 3060 3179 VWFD 10 (PARTIAL).

```

```

FT DOMAIN 3180 3299 WFFD 11 (PARTIAL).
FT DOMAIN 3300 3416 WFFD 12 (PARTIAL).
FT DOMAIN 3417 3536 WFFD 13 (PARTIAL).
FT DOMAIN 3537 3656 WFFD 14 (PARTIAL).
FT DOMAIN 3657 3776 WFFD 15 (PARTIAL).
FT DOMAIN 3777 3892 WFFD 16 (PARTIAL).
FT DOMAIN 3893 4029 WFFD 17 (PARTIAL).
FT DOMAIN 4029 4148 WFFD 18 (PARTIAL).
FT DOMAIN 4149 4263 WFFD 19 (PARTIAL).
FT DOMAIN 4264 4383 WFFD 20 (PARTIAL).
FT DOMAIN 4384 4503 WFFD 21 (PARTIAL).
FT DOMAIN 4504 4623 WFFD 22 (PARTIAL).
FT DOMAIN 4624 4743 WFFD 23 (PARTIAL).
FT DOMAIN 4744 4863 WFFD 24 (PARTIAL).
FT DOMAIN 4864 5295 WFFD 25.
FT DISULFID 5263 5274 BY SIMILARITY.
FT DISULFID 5268 5283 BY SIMILARITY.
FT DISULFID 5285 5294 BY SIMILARITY.
FT CARBOHYD 339 339 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 499 499 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1216 1216 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1239 1239 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1314 1314 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1814 1814 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1908 1908 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1933 1933 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 2028 2028 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 2111 2111 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 2142 2142 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 2332 2332 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 2533 2533 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 2575 2575 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 2692 2692 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 2812 2812 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 3052 3052 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 3065 3065 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 3144 3144 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 3172 3172 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 3288 3288 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 3292 3292 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 3782 3782 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 4005 4005 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 4136 4136 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 4243 4243 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 4254 4254 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 4335 4335 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 4376 4376 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 4586 4586 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 5136 5136 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 5252 5252 N-LINKED (GLCNAC. .) (POTENTIAL).
FT SEQUENCE 5376 AA; 579908 MW; 0E44DB77DF2A2620 CRC64;

Query Match 8.18; Score 189; DB 1; Length 5376;
Best Local Similarity 21.28; Pred. No. 4e-05;
Matches 111; Conservative 56; Mismatches 204; Indels 152; Gaps 28;

QY 23 CPDGTQAGLTGVAADLGTVC-----NCRPNFYNGGAAQGEANGNOPEA 69
DB 2437 CPPIGYCKN--SNDGSSN--CVKLSLOCPAHSKFTDCLPCHPSCSDPDGHCEGISTNA 2491

QY 70 ANNAARG-ICVP-----CQINRVGSV-----TNAG-----DLATLATQ 101
DB 2492 HSNKREGVCQGVYLRNDKCVLRIECGCQHTQGGFIPAGKNWTSRGCSDCMEGVIR 2551

QY 102 CST-QCPTGT---ALDDGVTDVDFRSAAC-----VKCKNFYNGSGSPQGEAPGVQVF 151
DB 2552 CQNFQCPSGTYCODIEDGTSNAC--ITLQCPAHSFTNCLPQPCSDPDGHCGSGSTTK 2610

QY 152 AAGAAAG-----VAAVTSQCPV-----CQ-----LNKNDSPATAGAANL 187
DB 2611 APSACQCGVCPEYVVLNKNKVPRIECGCKDAQGVLPADKIWNKGTCTQCAVGTGI 2670

QY 188 ATQCSN-OCPTGTVLDDGVTLVFNTSATLCVK-----CRPNFYNGSGSQGE 233
DB 2671 --HCRDFQCPSGTYCKD-----IKDDASNCTEIILOCPDHSLYTHCLPSCLLSCSDPDGL 2723

QY 234 AFGVQVFAAGAAAGAAVAVTSQVPCQINKNDSPATAGANLATQCTGPTGTATQDG 293
DB 2724 CRGTSPEAPSTCKEGCV-----CDPDYVLSND-----KCVLRIECGCKDAQGVLPAG 2771

QY 294 VTLVFSNSTQCSQCIANYFFGNF-----EAGKSQCLCKPVSRTTTPAHAFGNT 342
DB 2772 KTWI-NRGCTQSCMGAIOCNFKCPSEAYCOYMEDGNSNCTSIPLQ--CPAH-----2823

QY 343 ATQATQCLTTC-PAGTVLD---DGTSTNFVASATECTKCSAGFF-----ASK 385
DB 2824 -SHYTNCPLTPCQSCSDPDGHCEGSGSTRAPACKEGCVCEPDYVVLNKNKVPRIECGCKD 2882

QY 386 TTGFTAGTDTCTECTCKLITSGATAYAAATQKVQCASTTFEAK 428
DB 2883 TQGVLPADK-TWNRGCTQSCTRCGGAIOQCKYHCSSGTYCK 2924

RESULT 6
C170_GIALA STANDARD; PRT; 328 AA.
ID C170_GIALA AC P15799;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE Surface antigen Crp170 (Fragment).
OS Giardia lamblia (Giardia intestinalis).
OC Eukaryota; Diplomonadida; Hexamitidae; Giardia.
OX NCBI_Taxid=5741;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 30957 / WB;
RX MEDLINE=88089405; PubMed=3335828;
RA Adam R.D.; Aggarwal A.; Lal A.A.; de la Cruz V.F.; McCutchan T.;
RA Nash T.E.;
RT "Antigenic variation of a cysteine-rich protein in Giardia lamblia.";
RL J. Exp. Med. 167:109-118(1988).
CC -!- MISCELLANEOUS: CYSTEINE-RICH, ANTIGENICALLY VARIANT SURFACE
CC PROTEIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X06741; CAA29916.1; -.
DR PIR; S00530; S00530.
DR InterPro; IPR002174; Furin-like.
DR SMART; SM00261; FU; 3.
KW Repeat; Antigen.
FT NON_TER 1 1
FT REPEAT 38 102
FT REPEAT 103 167
FT NON_TER 328 328
FT SEQUENCE 328 AA; 33438 MW; 373A697A30EDCA21 CRC64;

Query Match 8.0%; Score 187; DB 1; Length 328;
Best Local Similarity 24.38; Pred. No. 3.9e-06;
Matches 102; Conservative 30; Mismatches 155; Indels 132; Gaps 25;

QY 20 AVPCPDGTQTAG-LTVDGAADLGTVCNCRPNFYNGGAAQGEANGNOPTAANNAARGIC 78
DB 13 AVDC-----QGSAGYTTDDSVSDAKECKK-----NAPCTACAGTADKC 51

QY 79 VPCQINRVG---SVTNAGD-LATLATQCTGPTGATLDDGVTDVDFRSAQCVKCPNF 134
DB 52 TKCDANGAARYLKKNPDPDTGTCVSAVDCQGSAGYTTDDSVSD-----AKECKCN--- 103
```


QY 135 YYNGSGPGEAPGVQVFAAGAAAGVAAVTSQVPCQLN-----KNDSPATAGAQAANLA 188
 Db 104 -----APCTACAGTA---DKTKCDANGAAPLYLTKTNSDPTGTGCVSA 143
 QY 189 TQCSNQCPGTGVLDDGVTLVFNTSATLCVCRPNFYNGSGSPOGEAPGVQVFAAGAAAG 248
 Db 144 VDC--QGSAGYTTDSV-----SDAKECKC-----AGQCKPN-----TAG 177
 QY 249 VAAVTSQVPCQLNKNDSPATAGAQAANLATQCSST-----QCPGTGT-----AIQDGYT-L 296
 Db 178 -----TQCFSC-----SDANCERCQDNDVCARCSTGAPPENGKCPAATPGCHSSCDGCTEN 228
 QY 297 VFSNSTQCSQCIANYEFNGFAGKS-QCL---KCPVSKTTPAHAPGNTATQATQCLTT 352
 Db 229 AMTNOADCKTGCKRGYLLKPESSAAGSGACLTAECTSDKT---HTREKAGSGKMGCLS 285
 QY 353 CPAGTVLDDGTSNFINVASATECTKCSAGFFASKT-TGFTAGTDTCTECT-KLITSGATA 409
 Db 286 CSDAT-----HGITGCKKC-----ALKTLGAEABSTVVCSECTDKRLTPSGNA 328
 RESULT 7
 TS11_GIALA
 ID TS11_GIALA STANDARD; PRT; 667 AA.
 AC Q03185;
 DT 01-JUN-1994 (Rel. 29, Created)
 DT 01-JUN-1994 (Rel. 29, Last sequence update)
 DT 15-JUL-1998 (Rel. 36, Last annotation update)
 DE Major surface trophozoite antigen 11 precursor.
 GN TS11.
 OS Giardia lamblia (Giardia intestinalis).
 OC Eukaryota; Diplomonadida; Hexamitidae; Giardinae; Giardia.
 OX NCBI_TaxID=5741;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Isolate AD-1;
 RC MEDLINE=93241215; PubMed=8479449;
 RA Ey P.L., Khanna K., Manning P.A., Mayrhofer G.;
 RT "A gene encoding a 69-kilodalton major surface protein of Giardia
 intestinalis trophozoites";
 RL Mol. Biochem. Parasitol. 58:247-258(1993).
 CC -! SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN. ANCHORED TO THE
 CC PLASMA MEMBRANE.
 CC -! DOMAIN: CONTAINS 27 REPEATS OF A CXXC MOTIF.
 CC -! SIMILARITY: BELONGS TO THE GIARDIA VARIANT SURFACE PROTEIN FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; M95814; AAA02687.1; .
 DR FIR; A48579; A48579.
 DR HSP; P02468; 1KLO.
 DR InterPro; IPR000561; EGF-like.
 DR InterPro; IPR002174; Furin-like.
 DR InterPro; IPR005127; Giardia_VSP.
 DR Pfam; PF03302; VSP; 1.
 DR SMART; SM00181; EGF; 3.
 DR SMART; SM00261; FU; 5.
 KW Antigen; Repeat; Transmembrane; Signal.
 FT SIGNAL 1 17
 FT CHAIN 18 667 MAJOR SURFACE TROPHOZOITE ANTIGEN 11.
 FT DOMAIN 18 633 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 634 662 POTENTIAL.
 FT DOMAIN 663 667 CYTOPLASMIC (POTENTIAL).
 FT CARBOHYD 591 591 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 630 630 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 667 AA; 68475 MW; 1DD95727032328BD CRC64;

Query Match 7.8%; Score 182; DB 1; Length 667;
 Best Local Similarity 22.2%; Pred No. 1.7e-05;
 Matches 120; Conservative 42; Mismatches 170; Indels 208; Gaps 32;
 QY 7 LLIISLFTNEURAVPCPDGTQT-----QAGLTDV-----GAADL-GTCVNCRPNFY 52
 Db 1 MLLAIFYFVISTITAKTCTTQTCTCEAEKCEMVGETEICTRCOTGKVPIDGKCVDATAN-- 58
 QY 53 YNGGAAGGANGNPPFAANNAARG--ICVPCQINRVGSVT-----NAGDLATLATQC 102
 Db 59 ANCKNAGSGGDANQVCKMSSVPGNTLCITVSPDGVSVAANEYFVPPNADATHDSVVSVC 118
 QY 103 STQCPGTGTLDDGVTDFDRSAAQCVCCKPNFYNGSGSPOGEAPGVQVFAAGAAAGVAA 162
 Db 119 SEETPHLANNQYIGV-----AGCATC-----SAP--KAPGEDNTPRAAT----- 157
 QY 163 VTSQVPCQLNKNDSPATAGAQAANLATQCSNOCPTGTVDLDDGVTLVFNTSAT-----LCV 217
 Db 158 ----CTKANGFLHTPSEG-----LSCEETCPPEG-----YFGHTATAESKKTCK 198
 QY 218 KCRPNFYNGSGSPOGEAPGVQVFAAGAAAGVAAVTSQVPCQINKNKNDSPATAGAQAANLA 277
 Db 199 SC-----TGGG--SEAPNVK-----GI-----GDCLKMYNE-----ASGNL 229
 QY 278 T--QCSTQ-----CPTGTATQDQGVTLVFSNSTQ-- 304
 Db 230 TCEKCSAOKKPSLDKTKTSCNDCTGQNCACFCSSGGDCGCDG--FILDGQNCVKSDCKTEN 288
 QY 305 -----CSQCIANYFFN-----GNFEAG-----KSQCLKCPVS--K 332
 Db 289 CKACTNPKAANEVCECTISTHLLTPTSOCVQYCOALGNFYAGTNADNKKKCKECTVANCK 348
 QY 333 T-----TPAHAPGNT--ATQATQCLTTCPTAGTVL---DDGTSTN 366
 Db 349 TCNDQGCQTCNDGFGYKNGDACSPCHESCKTCSAGTASDC-TECTPGKALAYGNDGTK-- 405
 QY 367 FVASATECTKCSAGFFASKTTGFTA-GTDTCTECTKKLTSGATAKVAEATQKVOCASSTT 425
 Db 406 --GTCGEGCTTGQSGGACKTCLTIDGASYCSECDTQ-----NEYPQNGICTSTT 453
 RESULT 8
 LMA5_HUMAN
 ID LMA5_HUMAN STANDARD; PRT; 3695 AA.
 AC O15230; Q9HIP1; Q8WZA7;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 15-JUN-2002 (Rel. 41, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Laminin alpha-5 chain precursor.
 GN LMA5 OR KIAA0533 OR KIAA1907.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA MEDLINE=21638749; PubMed=11780052;
 RA Deloukas P., Matthews L.H., Ashurst J., Burton J., Gilbert J.G.R.,
 RA Jones M., Stavrides G., Almeida J.P., Babbage A.K., Baguley C.L.,
 RA Bailey J., Barlow K.F., Bates K.N., Beard L.M., Beare D.M.,
 RA Beasley O.P., Bird C.P., Blakey S.E., Bridgeman A.M., Brown A.J.,
 RA Buck D., Burrill W.D., Butler A.P., Carder C., Carter N.P.,
 RA Chapman J.C., Clamp M., Clark A.P., Clark L.N., Clark S.Y., Clee C.M.,
 RA Clegg S., Cobley V.E., Collier R.E., Connor R.E., Dunn N.R.,
 RA Coulson A., Coville G.J., Deadman R., Dhani P.D., Durr M.,
 RA Ellington A.G., Frankland J.A., Fraser A., French L., Garner P.,
 RA Grafham D.V., Griffiths J.C., Griffiths M.N.D., Gwilliam R., Hall R.E.,
 RA Hammond S., Harley J.L., Heath P.D., Ho S., Holden J.L., Howden P.J.,
 RA Huckle E., Hunt A.R., Hunt S.E., Jekosch K., Johnson C.M., Johnson D.,
 RA Kay M.P., Kimberley A.M., King A., Knights K., Laird G.K., Lawlor S.,
 RA Lechvasiaho M.H., Leversha M.A., Lloyd C., Lloyd D.M., Lovell J.D.,
 RA Marsh V.L., Martin S.L., McConnachie L.J., McLay K., McMurray A.A.,

FT DISULFID 408 426 BY SIMILARITY.
FT DISULFID 429 440 BY SIMILARITY.
FT DISULFID 431 447 BY SIMILARITY.
FT DISULFID 449 458 BY SIMILARITY.
FT DISULFID 461 471 BY SIMILARITY.
FT DISULFID 494 506 BY SIMILARITY.
FT DISULFID 496 515 BY SIMILARITY.
FT DISULFID 517 526 BY SIMILARITY.
FT DISULFID 529 538 BY SIMILARITY.
FT DISULFID 541 553 BY SIMILARITY.
FT DISULFID 543 560 BY SIMILARITY.
FT DISULFID 562 571 BY SIMILARITY.
FT DISULFID 574 584 BY SIMILARITY.
FT DISULFID 587 599 BY SIMILARITY.
FT DISULFID 589 605 BY SIMILARITY.
FT DISULFID 607 616 BY SIMILARITY.
FT DISULFID 619 629 BY SIMILARITY.
FT DISULFID 632 644 BY SIMILARITY.
FT DISULFID 634 650 BY SIMILARITY.
FT DISULFID 652 661 BY SIMILARITY.
FT DISULFID 664 674 BY SIMILARITY.
FT DISULFID 677 689 BY SIMILARITY.
FT DISULFID 679 696 BY SIMILARITY.
FT DISULFID 698 707 BY SIMILARITY.
FT DISULFID 710 720 BY SIMILARITY.

Query Match 7.8%; Score 182; DB 1; Length 3695;
Best Local Similarity 21.8%; Pred No. 8.5e-05;
Matches 97; Conservative 38; Mismatches 171; Indels 138; Gaps 25;

QY 43 TCVNCRNFY-----YNGGAQGEANGNPFANNAARGICVPCQIN-----RVG 87
Db 1841 SCOECPAGFYRVKGLFLGRCVPCQCHGHSRCLPGS--GVCVDQHTEGAHCRCQAG 1898
QY 88 SVTNAGDLATATQCTQCP-----TGTALEDGVDVDFRSDAAQCVCKPNFYNG 138
Db 1899 FVSRDDPSAPVCSC--PCPLSVPSNNFAEGCVLRGRT-----QCL-CKPG--YAG 1945
QY 139 GSPQGEAPGVQVFAAGAAAGVAAYTSQVPCQKNDSPATAGAAQANLATQCSNQCPGT 198
Db 1946 ASCERCAPG---FFGNPLVLG-----SSCPDCSGNDPNLLSDCDPLGACRGC--- 1994
QY 199 TVLDDGVTLVFNTSATLCKVRPNFYNGSGPQGEAPGVQVFAAGAAAG-----VA 250
Db 1995 -----LRHTTGRCIECAPGYGNALLP-GNCTRCDCCTPGCTEACDPHSGHCLCKA 2044
QY 251 AVTSQ-----CVPCQKNDSPATAGAAQAN-LATQCTQCTPGTAIQD 292
Db 2045 GVTGRCRDCQEGHFGNCGGCRFCAC-----GPAEGSECHPQSGQC--HCRPGT--- 2094
QY 293 GVTLVFSNSTOCOCIANYFNGNFEAGKSQLCKPVSRTTPAHAGNTATQATQCLTT 352
Db 2095 -----NGPQCRECAPGYW--GLPEQGCRRC-QCPGGRCDP-----HTGRC--N 2132
QY 353 CPAGTVLDGDTSTNEVASATECTKCSAGF---FASKTTGFTAGTDTCTECTKKL-----T 404
Db 2133 CPFG-----LSGERCDTCSQHQHVPVPGPGVSHICEVCHDCVLLDLDL 2180
QY 405 SCATAKYVAETAKVQCASTTFAX 428
Db 2181 AGALLPAIHEQLRGINASSMAWAR 2204

RESULT 9
LMA1_HUMAN
ID LMA1_HUMAN STANDARD; PRT: 3075 AA.
AC P25391;
DT 01-MAY-1992 (Rel. 22, Created)
DT 01-MAY-1992 (Rel. 22, Last sequence update)
DE 15-JUN-2002 (Rel. 41, Last annotation update)
DE Laminin alpha-1 chain precursor (Laminin A chain).
GN LMA1 OR LAMA.
OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91333420; PubMed=1714537;
RA Haaparanta T., Uitto J., Ruoslahti E., Engvall E.;
FT "Molecular cloning of the cDNA encoding human laminin A chain.";
RL Matrix 11:151-160(1991).
RN [2]
RP SEQUENCE OF 1-2628 FROM N.A.
RX MEDLINE=91264789; PubMed=2049067;
RA Nissinen M., Vuolteenaho R., Boot-Handford R., Kallunki P.,
RA Tryggvason K.;
FT "Primary structure of the human laminin A chain. Limited expression
in human tissues.";
RL Biochem. J. 276:369-379(1991).
RN [3]
RP SEQUENCE OF 2397-3072 FROM N.A.
RX MEDLINE=89280632; PubMed=2733383;
RA Olsen D., Nagayoshi T., Fazio M., Peltonen J., Jaakkola S.,
RA Sanborn D., Sasaki T., Kuivaniemi H., Chu M.L., Deutzmann R.,
RA Timpl R., Uitto J.;
FT "Human laminin: cloning and sequence analysis of cDNAs encoding A, B1
and B2 chains, and expression of the corresponding genes in human
skin and cultured cells.";
RL Lab. Invest. 60:772-782(1989).
CC -|- FUNCTION: Binding to cells via a high affinity receptor, laminin
is thought to mediate the attachment, migration, and organization
of cells into tissues during embryonic development by interacting
with other extracellular matrix components.
CC -|- SUBUNIT: Laminin is a complex glycoprotein, consisting of three
different polypeptide chains (alpha, beta, gamma), which are bound
to each other by disulfide bonds into a cross-shaped molecule
comprising one long and three short arms with globules at each
end.
CC THE ALPHA-1 CHAIN IS A SUBUNIT OF LAMININ-1 (BHS LAMININ) AND
LAMININ-3 (S-LAMININ).
CC -|- SUBCELLULAR LOCATION: Extracellular.
CC -|- TISSUE SPECIFICITY: FOUND IN THE BASEMENT MEMBRANES (MAJOR
COMPONENT).
CC -|- DOMAIN: THE ALPHA-HELICAL DOMAINS I AND II ARE THOUGHT TO INTERACT
WITH OTHER LAMININ CHAINS TO FORM A COILED COIL STRUCTURE.
CC -|- DOMAIN: DOMAINS VI, IV AND G ARE GLOBULAR.
CC -|- SIMILARITY: CONTAINS 1 LAMININ N-TERMINAL DOMAIN (DOMAIN VI).
CC -|- SIMILARITY: CONTAINS 17 LAMININ EGF-LIKE DOMAINS.
CC -|- SIMILARITY: CONTAINS 2 LAMININ DOMAINS IV.
CC -|- SIMILARITY: CONTAINS 5 LAMININ G-LIKE DOMAINS.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL Outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X58531; CAA41418.1; -;
CC PIR; S14458; S14458.
CC HSSP; Q60675; LQ00.
CC Genew; HGNC:6481; LAMA1.
CC MIM; 150320; -;
CC InterPro; IPR000561; EGF-like.
CC InterPro; IPR01886; LamNT.
CC InterPro; IPR000034; Laminin_B.
CC InterPro; IPR002049; Laminin_G.
CC Pfam; PF00052; laminin_B; 2.
CC Pfam; PF00053; laminin_EGF; 15.
CC Pfam; PF00054; laminin_G; 5.
CC Pfam; PF00055; laminin_Nterm; 1.
CC PRINTS; PR00011; EGF/LAMININ.
CC ProDom; PD002082; LamNT; 1.

[illegible]

[4]
X-RAY CRYSTALLOGRAPHY (2.35 ANGSTROMS) OF 2932-3106.
MEDLINE=20085745; PubMed=10619025;
Hohenester E., Tisi D., Ralts J.F., Timpl R.;
"The crystal structure of a laminin G-like module reveals the
molecular basis of alpha-dystroglycan binding to laminins, perlecan,
and agrin.";
Mol. Cell 4:783-792(1999).
-!- FUNCTION: Binding to cells via a high affinity receptor, laminin
is thought to mediate the attachment, migration, and organization
of cells into tissues during embryonic development by interacting
with other extracellular matrix components.
-!- SUBUNIT: Laminin is a complex glycoprotein, consisting of three
different polypeptide chains (alpha, beta, gamma), which are bound
to each other by disulfide bonds into a cross-shaped molecule
Comprising one long and three short arms with globules at each
end.
THE ALPHA-2 CHAIN IS A SUBUNIT OF LAMININ-2 (MEROSIN) AND LAMININ-
4 (S-MEROSIN).
-!- SUBCELLULAR LOCATION: Extracellular.
-!- TISSUE SPECIFICITY: FOUND IN THE BASEMENT MEMBRANES (MAJOR
COMPONENT).
-!- DOMAIN: THE ALPHA-HELICAL DOMAINS I AND II ARE THOUGHT TO INTERACT
WITH OTHER LAMININ CHAINS TO FORM A COILED COIL STRUCTURE.
-!- DOMAIN: DOMAINS VI, IV AND G ARE GLOBULAR.
-!- DISEASE: DEFECTS IN LAMA2 ARE A CAUSE OF MURINE MUSCULAR DYSTROPHY
(DY2J).
-!- SIMILARITY: CONTAINS 1 LAMININ N-TERMINAL DOMAIN (DOMAIN VI).
-!- SIMILARITY: CONTAINS 17 LAMININ EGF-LIKE DOMAINS.
-!- SIMILARITY: CONTAINS 2 LAMININ DOMAINS IV.
-!- SIMILARITY: CONTAINS 5 LAMININ G-LIKE DOMAINS.

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).

CC EMBL; U12147; AAC52165.1; -;
DR EMBL; X69869; CAA49502.1; -;
DR EMBL; S75315; AAB33573.1; -;
DR PDB; 1QU0; 03-DEC-99.
DR MGD; MGI:99912; Lama2.
DR InterPro; IPRO00561; EGF-like.
DR InterPro; IPRO01886; LamNT.
DR InterPro; IPRO00034; Laminin_B.
DR InterPro; IPRO02049; Laminin_EGF.
DR InterPro; IPRO01791; Laminin_G.
DR Pfam; PF00052; laminin_B; 2.
DR Pfam; PF00053; laminin_EGF; 15.
DR Pfam; PF00054; laminin_G; 5.
DR Pfam; PF00055; laminin_Nterm; 1.
DR PRINTS; PR00011; EGFLAMINRM.
DR ProDom; PD002082; LamNT; 1.
DR ProDom; PD03031; Laminin_B; 2.
DR SMART; SM00180; EGF_Lam; 15.
DR SMART; SM00001; EGF_like; 1.
DR SMART; SM00281; Lamb; 2.
DR SMART; SM00282; Lang; 5.
DR SMART; SM00136; LamNT; 1.
DR PROSITE; PS00022; EGF_1; 11.
DR PROSITE; PS01186; EGF_2; 3.
DR PROSITE; PS01248; LAMININ_TYPE_EGF; 14.
DR PROSITE; PS50025; IAM_G_DOMAIN; 5.
KW Glycoprotein. Basement membrane; Extracellular matrix; Coiled coil;
Laminin EGF-like domain; Cell adhesion; Repeat; Signal; 3D structure.
FT SIGNAL 1 22 POTENTIAL.
FT CHAIN 23 3106 LAMININ ALPHA-2 CHAIN.
FT DOMAIN 23 282 LAMININ N-TERMINAL (DOMAIN VI).
FT DOMAIN 283 339 LAMININ EGF-LIKE 1.
FT DOMAIN 340 409 LAMININ EGF-LIKE 2.

RESULT 13
 LMG3_HUMAN
 ID LMG3_HUMAN STANDARD; PRT; 1587 AA.
 AC Q9Y6N6;
 DT 15-JUN-2002 (Rel. 41, Created)
 DT 15-JUN-2002 (Rel. 41, Last sequence update)
 DE Laminin gamma-3 chain precursor (Laminin 12 gamma 3).
 GN LMG3.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Placenta;
 RX MEDLINE=99242614; PubMed=10225960;
 RA Koch M., Olson P.F., Albus A., Jin W., Hunter D.D., Brunken W.J.,
 RA Burgeson R.E., Champlaud M.F.;
 RT "Characterization and expression of the laminin gamma3 chain: a novel,
 RT non-basement membrane-associated, laminin chain.";
 RL J. Cell Biol. 145:605-618(1999).
 CC -1- FUNCTION: Binding to cells via a high affinity receptor, laminin
 CC is thought to mediate the attachment, migration, and organization
 CC of cells into tissues during embryonic development by interacting
 CC with other extracellular matrix components.
 CC -1- SUBUNIT: Laminin is a complex glycoprotein, consisting of three
 CC different polypeptide chains (alpha, beta, gamma), which are bound
 CC to each other by disulfide bonds into a cross-shaped molecule
 CC comprising one long and three short arms with globules at each
 CC end.
 CC THE GAMMA-3 CHAIN IS A SUBUNIT OF LAMININ-12.
 CC -1- SUBCELLULAR LOCATION: Extracellular.
 CC -1- TISSUE SPECIFICITY: Broadly expressed in: skin, heart, lung, and
 CC the reproductive tracts.
 CC -1- DOMAIN: THE ALPHA-HELICAL DOMAINS I AND II ARE THOUGHT TO INTERACT
 CC WITH OTHER LAMININ CHAINS TO FORM A COILED COIL STRUCTURE.
 CC -1- DOMAIN: DOMAIN IV IS GLOBULAR.
 CC -1- SIMILARITY: CONTAINS 1 LAMININ N-TERMINAL DOMAIN (DOMAIN VI).
 CC -1- SIMILARITY: CONTAINS 11 LAMININ EGF-LIKE DOMAINS.
 CC -1- SIMILARITY: CONTAINS 1 LAMININ DOMAIN IV.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL: AF041835; AAD36991.1; -;
 CC HSSP: P02468; 1TLE.
 CC Genew: HGNC:6494; LMG3.
 CC MIM: 604349; -;
 CC InterPro: IPR000561; EGF-like.
 CC InterPro: IPR001886; LamNT.
 CC InterPro: IPR000034; Laminin_B.
 CC InterPro: IPR002049; Laminin_EGF.
 CC Pfam: PF00052; laminin_B; 1.
 CC Pfam: PF00053; laminin_EGF; 10.
 CC Pfam: PF00055; laminin_Nterm; 1.
 CC PRINTS: P00011; EGF_LAMININ.
 CC ProDom: P0002082; LamNT; 1.
 CC ProDom: P0003031; Laminin_B; 1.
 CC SMART: SM00180; EGF_Lam; 10.
 CC SMART: SM00001; EGF_Like; 1.
 CC SMART: SM00281; Lam5; 1.
 CC SMART: SM00136; LamNT; 1.
 CC PROSITE: PS00022; EGF_1; 7.
 CC PROSITE: PS01186; EGF_2; 2.
 CC PROSITE: PS01248; LAMININ_TYPE_EGF; 10.
 CC Glycoprotein; Basement membrane; Extracellular matrix; Coiled coil;
 KW Laminin EGF-like domain; Cell adhesion; Repeat; Signal.

FT	SIGNAL	1	19	POTENTIAL
FT CHAIN	20	1587		LAMININ GAMMA-3 CHAIN.
FT DOMAIN	20	270		LAMININ N-TERMINAL (DOMAIN VI).
FT DOMAIN	271	326		LAMININ EGF-LIKE 1.
FT DOMAIN	327	382		LAMININ EGF-LIKE 2.
FT DOMAIN	383	429		LAMININ EGF-LIKE 3.
FT DOMAIN	430	479		LAMININ EGF-LIKE 4.
FT DOMAIN	480	489		LAMININ EGF-LIKE 5 (N-TERMINAL).
FT DOMAIN	490	672		LAMININ DOMAIN IV.
FT DOMAIN	673	706		LAMININ EGF-LIKE 5 (C-TERMINAL).
FT DOMAIN	707	754		LAMININ EGF-LIKE 6.
FT DOMAIN	755	809		LAMININ EGF-LIKE 7.
FT DOMAIN	810	865		LAMININ EGF-LIKE 8.
FT DOMAIN	866	916		LAMININ EGF-LIKE 9.
FT DOMAIN	917	964		LAMININ EGF-LIKE 10.
FT DOMAIN	965	1013		LAMININ EGF-LIKE 11.
FT DOMAIN	1014	1587		DOMAIN II AND I.
FT DOMAIN	1071	1141		COILED COIL (POTENTIAL).
FT DOMAIN	1200	1229		COILED COIL (POTENTIAL).
FT DOMAIN	1424	1504		COILED COIL (POTENTIAL).
FT DOMAIN	1535	1579		COILED COIL (POTENTIAL).
FT SITE	1059	1061		CELL ATTACHMENT SITE (POTENTIAL).
FT CARBOHYD	87	87		N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD	119	119		N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD	295	295		N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD	328	328		N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD	631	631		N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD	837	837		N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD	980	980		N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD	1185	1185		N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD	1518	1518		N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD	1587	1587		N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE	1587 AA;	172051 MW;	3CB6E09B5F203319 CRC64;	

Query Match 7.5%; Score 175; DB 1; Length 1587;
 Best Local Similarity 19.7%; Pred. No. 0.00011;
 Matches 89; Conservative 43; Mismatches 155; Indels 164; Gaps 21;

QY	13	LFTELRAVPCPGDTQTQAGLTVDVGAADLG----	TCVNCRPNFYNGGAQGEANGNOPF	68
DB	649	VFLTEVRLTSARPLSPASWVEICSCPTGYTGQFCESCAPGY-----	KREMPQGGPY	701
QY	69	AANNAARGICVPCQINRVGSVTNAGDLATLATQCS--	TQCPGTALDDGVTVDVDRSAQ	126
DB	702	AS-----CVPTCQHQGTGTC-----	DPNTGICVCSHTEGFS-	732
QY	127	CVKCKPNFYNGGSP--QGEAPGVQVFAAGAAAA--		158
DB	733	CERCLPGFY---GNPPAGQADDQCPCPGQSACTTIPESGEVCTHCPPGGRRCVC	789	
QY	159	-----GVAAVTSQCVPCQINKNDSPTAGAAANLATQCSNQCPTGTVLDGVTLVF	209	
DB	790	DGFFGDPGLGLFQHPQCHQCQCSGNVDPNVNGCDPLSGHC-----	LRCLH	836
QY	210	NTSATLCVKCRPNFYNGSGSGEAPGVQVFAAGAAAGAAVTSQCVPCQINKNDSPT	269	
DB	837	NTGDCHEHCQEGFY---GSALAPRP-----	ADKCMPCSCHPQGSVSE	876
QY	270	AGAAANLATQCSCTQCTGTGTATQIDQVTLVFSNDSSTQCSQCIANYFNGFAGKSOCLKCP	329	
DB	877	QMPCDPVTGQCS--CLPHVTARD-----	CSRCYPGFF---DLQPRG-CRSC-	917
QY	330	VSKTTPAHAPGNTATQ-----ATQCLTTCAGTVDGTTSTNFVASATECTKCSAGFASK	385	
DB	918	-----KCHPLGSQEDQCHPKTQC--TCRPG-----	VTGQACDRCOLGFFGSS	958
QY	386	TTGFTAGTDTCTECTKLTSGATAKVVAEAT	416	
DB	959	IKG-----CRACRCSPLGAASACQCHYNGT	982	

RESULT 14
 LMB1_DRONE
 ID LMB1_DRONE STANDARD; PRT; 1790 AA.

AC P11046; Q26328; Q9XZT4; Q9VLW6;
 DT 01-JUL-1989 (Rel. 11, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Laminin beta-1 chain precursor (Laminin B1 chain).
 GN LAMB1 OR LAMB1 OR C67123.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;
 OC Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
 OC Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Canton-S;
 RX MEDLINE=94000382; PubMed=8397815;
 RA Gow C.-H., Chang H.-Y., Lih C.-J., Chang T.-W., Hui C.-F.;
 FT "Analysis of the Drosophila gene for the laminin B1 chain.";
 RL DNA Cell Biol. 12:573-587(1993).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=88210471; PubMed=3365769;
 RA Montell D.J., Goodman C.S.;
 RT "Drosophila substrate adhesion molecule: sequence of laminin B1 chain
 reveals domains of homology with mouse.";
 RL Cell 53:463-473(1988).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Berkely;
 RX MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Celnikier S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scherer S.E., Richards S., Ashburner M., Henderson S.N.,
 RA George R.A., Lewis S.E., Yandell M.D., Zhang Q., Chen L.X.,
 RA Sutton G.G., Wortman J.R., Blakes R.G., Champe M., Pfeiffer B.D.,
 RA Brandon R.C., Rogers J.-H.C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA Wan K.H., Doyle C., Aghayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
 RA Abell J.F., Aghayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
 RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,
 RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablo B., Delcher A., Deng Z., Drenth J., Dew I., Dietz S.M.,
 RA Dodson K.J., Douc L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
 RA Fostler A., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Swirski R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of Drosophila melanogaster.";
 RL Science 287:2185-2195(2000).
 CC -1- FUNCTION: Binding to cells via a high affinity receptor, laminin
 is thought to mediate the attachment, migration, and organization
 of cells into tissues during embryonic development by interacting
 with other extracellular matrix components.
 CC -1- SUBUNIT: Laminin is a complex glycoprotein, consisting of three
 different polypeptide chains (alpha, beta, gamma), which are bound

to each other by disulfide bonds into a cross-shaped molecule
 comprising one long and three short arms with globules at each
 end.
 -1- SUBCELLULAR LOCATION: Extracellular.
 -1- TISSUE SPECIFICITY: FOUND IN THE BASEMENT MEMBRANES (MAJOR
 COMPONENT).
 -1- DOMAIN: THE ALPHA-HELICAL DOMAINS I AND II ARE THOUGHT TO INTERACT
 WITH OTHER LAMININ CHAINS TO FORM A COILED COIL STRUCTURE.
 -1- DOMAIN: DOMAINS VI AND IV ARE GLOBULAR.
 -1- SIMILARITY: CONTAINS 1 LAMININ N-TERMINAL DOMAIN (DOMAIN VI).
 -1- SIMILARITY: CONTAINS 12.5 LAMININ EGF-LIKE DOMAINS.
 -1- SIMILARITY: CONTAINS 2 LAMININ DOMAINS IV.

 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).

 CC EMBL; M95811; AAD19752.1; -;
 CC EMBL; M19525; AAA28663.1; -;
 CC EMBL; AE003618; AAF52563.1; -;
 CC PIR; A28783; MMFFBI.
 CC HSP; P02468; ITLE.
 CC FlyBase; FBgn0002527; LanB1.
 CC InterPro; IPR000561; EGF-like.
 CC InterPro; IPR001886; LamNT.
 CC InterPro; IPR002049; Laminin_EGF.
 CC Pfam; PF00053; laminin_EGF; 13.
 CC PRINTS; PR00011; EGF_LAMININ.
 CC PRODOM; PD002082; LamNT; 1.
 CC SMART; SM00180; EGF_Lam; 11.
 CC SMART; SM00001; EGF_Like; 1.
 CC SMART; SM00136; LamNT; 1.
 CC PROSITE; PS00022; EGF_1; 10.
 CC PROSITE; PS00186; EGF_2; 2.
 CC PROSITE; PS01248; LAMININ_TYPE_EGF; 12.
 CC GlycoProtein; Basement membrane; Extracellular matrix; Coiled coil;
 CC Laminin EGF-like domain; Cell adhesion; Repeat; Signal.
 CC SIGNAL 1 25
 CC CHAIN 26 1790 LAMININ BETA-1 CHAIN.
 CC DOMAIN 43 289 LAMININ N-TERMINAL (DOMAIN VI).
 CC DOMAIN 290 356 LAMININ EGF-LIKE 1.
 CC DOMAIN 357 419 LAMININ EGF-LIKE 2.
 CC DOMAIN 420 479 LAMININ EGF-LIKE 3.
 CC DOMAIN 480 530 LAMININ EGF-LIKE 4.
 CC DOMAIN 531 561 LAMININ EGF-LIKE 5 (INCOMPLETE).
 CC DOMAIN 562 789 LAMININ DOMAIN IV.
 CC DOMAIN 791 838 LAMININ EGF-LIKE 6.
 CC DOMAIN 839 884 LAMININ EGF-LIKE 7.
 CC DOMAIN 885 934 LAMININ EGF-LIKE 8.
 CC DOMAIN 935 992 LAMININ EGF-LIKE 9.
 CC DOMAIN 993 1044 LAMININ EGF-LIKE 10.
 CC DOMAIN 1045 1095 LAMININ EGF-LIKE 11.
 CC DOMAIN 1096 1143 LAMININ EGF-LIKE 12.
 CC DOMAIN 1144 1190 LAMININ EGF-LIKE 13.
 CC DOMAIN 1191 1407 DOMAIN II.
 CC DOMAIN 1408 1434 DOMAIN ALPHA.
 CC DOMAIN 1435 1790 DOMAIN I.
 CC DOMAIN 1757 1407 COILED COIL (POTENTIAL).
 CC DOMAIN 1455 1507 COILED COIL (POTENTIAL).
 CC DOMAIN 1542 1563 COILED COIL (POTENTIAL).
 CC DOMAIN 1610 1764 COILED COIL (POTENTIAL).
 CC DOMAIN 290 299 BY SIMILARITY.
 CC DISULFID 292 320 BY SIMILARITY.
 CC DISULFID 322 331 BY SIMILARITY.
 CC DISULFID 334 354 BY SIMILARITY.
 CC DISULFID 357 366 BY SIMILARITY.
 CC DISULFID 359 384 BY SIMILARITY.
 CC DISULFID 387 396 BY SIMILARITY.

FT	DISULFID	399	417	BY SIMILARITY.
FT	DISULFID	420	433	BY SIMILARITY.
FT	DISULFID	422	448	BY SIMILARITY.
FT	DISULFID	450	459	BY SIMILARITY.
FT	DISULFID	462	477	BY SIMILARITY.
FT	DISULFID	480	493	BY SIMILARITY.
FT	DISULFID	482	500	BY SIMILARITY.
FT	DISULFID	502	511	BY SIMILARITY.
FT	DISULFID	514	528	BY SIMILARITY.
FT	DISULFID	791	803	BY SIMILARITY.
FT	DISULFID	793	810	BY SIMILARITY.
FT	DISULFID	812	821	BY SIMILARITY.
FT	DISULFID	824	836	BY SIMILARITY.
FT	DISULFID	839	851	BY SIMILARITY.
FT	DISULFID	841	858	BY SIMILARITY.
FT	DISULFID	860	869	BY SIMILARITY.
FT	DISULFID	872	882	BY SIMILARITY.
FT	DISULFID	885	894	BY SIMILARITY.
FT	DISULFID	887	901	BY SIMILARITY.
FT	DISULFID	904	913	BY SIMILARITY.
FT	DISULFID	916	932	BY SIMILARITY.
FT	DISULFID	935	951	BY SIMILARITY.
FT	DISULFID	937	962	BY SIMILARITY.
FT	DISULFID	964	973	BY SIMILARITY.
FT	DISULFID	976	990	BY SIMILARITY.
FT	DISULFID	993	1007	BY SIMILARITY.
FT	DISULFID	995	1014	BY SIMILARITY.
FT	DISULFID	1017	1026	BY SIMILARITY.
FT	DISULFID	1029	1042	BY SIMILARITY.
FT	DISULFID	1045	1059	BY SIMILARITY.
FT	DISULFID	1047	1066	BY SIMILARITY.
FT	DISULFID	1068	1077	BY SIMILARITY.
FT	DISULFID	1080	1093	BY SIMILARITY.
FT	DISULFID	1096	1108	BY SIMILARITY.
FT	DISULFID	1098	1115	BY SIMILARITY.
FT	DISULFID	1117	1126	BY SIMILARITY.
FT	DISULFID	1129	1141	BY SIMILARITY.
FT	DISULFID	1144	1156	BY SIMILARITY.
FT	DISULFID	1146	1163	BY SIMILARITY.
FT	DISULFID	1165	1174	BY SIMILARITY.
FT	DISULFID	1177	1188	BY SIMILARITY.
FT	DISULFID	1191	1191	INTERCHAIN (PROBABLE).
FT	DISULFID	1194	1194	INTERCHAIN (PROBABLE).
FT	DISULFID	1788	1788	INTERCHAIN (PROBABLE).
FT	SITE	643	645	CELL ATTACHMENT SITE (POTENTIAL).
FT	CARBOHYD	140	140	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	203	203	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	234	234	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	489	489	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	593	593	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	1053	1053	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	1248	1248	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	1303	1303	N-LINKED (GLCNAC. . .) (POTENTIAL).
Query Match 7.4%; Score 173; DB 1; Length 1790;				
Best Local Similarity 21.5%; Pred. No. 0.00017;				
Matches 97; Conservative 37; Mismatches 147; Indels 170; Gaps 23;				
Qy	10 IISLFINELRAVPC---	PDG	TQAGLTDVGAADLGTVCNRPFPYNGGAAQGEANGNQ	66
Db	779 ILSVFVHD--GASMCNCP	TGS-----	LSKVCESGGYV--QCKPNV--GRQCDQCAPGY	829
Qy	67 PFAANNAARGICVPCQ	INRVGSVTNAGDLATLATQCSTQCTGTALDDGVTDVDRSAQ	126	
Db	830 GFGPEG-----	CKACDCNSIGSKDYCDLITGQCQC-----	VNTYGR---E	868
Qy	127 CVKCKPFPYNGGSPQGE	APGVQVFAAGAAAGAAVTSQCVPQQLNKNDSPATAGAAQ	186	
Db	869 CNOCPQGYW-----	NFPECRVCQCNGHAATCDPIQGTCDICQ-----	DS-----	907
Qy	187 LATQCSNOCPGTGTVL	DDGVTLVFNTSATLCKVCRPNFYNG--GSPQGEAPGVQVFAAGA	244	
Db	908 -----	TTGYGDCSDLDGYGNPLFGSEIG-----	931	

```
DR EMBL; AF079520; AAD29851.1; -.
DR HSP; P02468; IKLO.
DR MGD; MGI:1344394; Lamc3.
DR InterPro; IPR000561; EGF-like.
DR InterPro; IPR001886; LamNT.
DR InterPro; IPR000034; Laminin_B.
DR InterPro; IPR002049; Laminin_EGF.
DR Pfam; PF00052; laminin_B_1.
DR Pfam; PF00053; laminin_EGF_10.
DR Pfam; PF00055; laminin_Nterm; 1.
DR PRINTS; PR00011; EGF_LAMININ.
DR ProDom; PD002082; LamNT; 1.
DR SMART; SM00180; EGF_Lam; 10.
DR SMART; SM00281; LamB; 1.
DR SMART; SM00136; LamNT; 1.
DR PROSITE; PS00022; EGF_1; UNKNOWN_8.
DR PROSITE; PS01186; EGF_2; 2.
DR PROSITE; PS01248; LAMININ_TYPE_EGF; 11.
DR Glycoprotein; Basement membrane; Extracellular matrix; Coiled coil;
KW Laminin EGF-like domain; Cell adhesion; Repeat; Signal.
FT SIGNAL 1 28
FT CHAIN 29 1581
FT DOMAIN 29 279 LAMININ GAMMA-3 CHAIN.
FT DOMAIN 280 335 LAMININ N-TERMINAL (DOMAIN VI).
FT DOMAIN 336 391 LAMININ EGF-LIKE 1.
FT DOMAIN 392 438 LAMININ EGF-LIKE 2.
FT DOMAIN 439 488 LAMININ EGF-LIKE 3.
FT DOMAIN 489 498 LAMININ EGF-LIKE 4.
FT DOMAIN 499 584 LAMININ EGF-LIKE 5 (N-TERMINAL).
FT DOMAIN 585 718 LAMININ DOMAIN IV.
FT DOMAIN 719 766 LAMININ EGF-LIKE 5 (C-TERMINAL).
FT DOMAIN 767 821 LAMININ EGF-LIKE 6.
FT DOMAIN 822 877 LAMININ EGF-LIKE 7.
FT DOMAIN 878 927 LAMININ EGF-LIKE 8.
FT DOMAIN 928 975 LAMININ EGF-LIKE 9.
FT DOMAIN 976 1024 LAMININ EGF-LIKE 10.
FT DOMAIN 1025 1581 LAMININ EGF-LIKE 11.
FT DOMAIN 1029 1046 LAMININ II AND I.
FT DOMAIN 1112 1153 COILED COIL (POTENTIAL).
FT DOMAIN 1208 1231 COILED COIL (POTENTIAL).
FT DOMAIN 1438 1468 COILED COIL (POTENTIAL).
FT DOMAIN 1510 1575 COILED COIL (POTENTIAL).
FT CARBOHYD 128 128 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 304 304 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 337 337 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 640 640 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 849 849 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 991 991 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1162 1162 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1196 1196 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1320 1320 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1514 1514 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CONFLICT 9 9 L -> F (IN REF. 2).
FT CONFLICT 190 190 P -> T (IN REF. 2).
FT CONFLICT 195 195 R -> K (IN REF. 2).
FT CONFLICT 221 221 G -> S (IN REF. 2).
FT CONFLICT 394 394 C -> R (IN REF. 2).
FT CONFLICT 471 471 C -> Y (IN REF. 2).
FT CONFLICT 1150 1150 L -> LDEPOLFSLKK (IN REF. 2).
FT CONFLICT 1387 1387 H -> Q (IN REF. 2).
FT CONFLICT 1438 1439 AS -> TI (IN REF. 2).
FT CONFLICT 1479 1479 V -> I (IN REF. 2).
SQ SEQUENCE 1581 AA; 172316 MW; 51DFAD1F95E6AE81 CRC64;

Query Match
Best Local Similarity 7.4%; Score 172.5; DB 1; Length 1581;
Matches 86; Conservative 40; Mismatches 131; Indels 159; Gaps 26;

QY 21 VPCPDGTQTQAGLTD--VGAADLG-----TCVNCRPNFYNGGAAQGEANGNQPFPAANN 72
Db 717 IPC---TCNQHGTCDPNTGICLGHHTGPGSCRCMPGFY--GNAFSGRADDCQP----- 766
QY 73 AARGICVPCQ--INRVGSVTNAGDLATLATQCSTQCPTGTALDDGVTDFDRSAOCVKCK 131
```

```
Db 767 -----CPCPGOSACATIPESGDVV-----CTHCPPG-----ORGRRCESCE 802
QY 132 PNFYNGSGSQGEAPGVQVFAAGAAAAGVAAVTSCVPCQLNKNKNDSPATAGAQAANLATQC 191
Db 803 DGFF---GDP-----LGLSGAPQPCRCQCQSGN-----VDLNAVNC 836
QY 192 ---SNQCPTGTVDGVTLVFNTSATLCVKCRPNFYNGSGPQGEAPGVQVFAAGAAAAG 248
Db 837 DPHSGHC-----LRCLYNTTGAHCEHCREGFY-----GSAYA- 868
QY 249 VAAVTSCVPCQIN-----KNDSPATAGAQAANLATQCSTQCPTGTALDQGVTLVESNS 301
Db 869 -TRPVDKCAPCSCDLRGSVSEKTCNPVTG-----QC-----VCLPYV-S 905
QY 302 STQCSOCIANYFFNGFNFAKGSQCLKCPVSKTTPAHAPGNTATQAT-QCLTTCPTAGTVLD 360
Db 906 GRDCSRCSGPFY---DLQSGRG-CQSC---KCHPLGSLKNKCHPKTGQC--PCRPQ---- 952
QY 361 DGTSTNFVASATECTKCSAGFFASKTTGTTAGTDTCTCTCKLTSGATAKYVAEAT 416
Db 953 -----VTGQACDRCLQGLFFGFSIKG-----CRDCRCSPLGAASSQCHENST 993
```

Search completed: February 11, 2003, 19:46:54
Job time : 20.7429 secs